P2.6 A BAYESIAN APPROACH TO RETRIEVE RAINDROP SIZE DISTRIBUTION FROM POLARIMETRIC RADAR DATA

Qing Cao and Guifu Zhang

University of Oklahoma, Norman, OK 73072

1. INTRODUCTION

The retrieval of raindrop size distribution (DSD) from polarimetric radar data (PRD) has the potential to improve the quantitative precipitation estimation (QPE), understanding precipitation microphysics and quantitative precipitation forecasts (QPF) (e.g., Brandes et al. 2004). Radar measurements and the retrieval model contain errors, which lead to errors in the retrieved DSDs. For example, if the retrieval model is based on the assumption of pure rain while the radar echoes sometimes includes snow, hail or non-hydrometeor effects, the retrieved DSD will be incorrect. In practice, however, it is difficult to know measurement and model errors. The Bayesian theory offers a promising method of optimizing the use of PRD for DSD retrievals. Using the prior information of measured DSDs, the Bayesian approach provides not only the mean values of DSD parameters, but also their standard deviations, which determine the reliability of retrievals. Two factors are required for the success of the Bayesian approach. The first one is the use of appropriate state parameters and conditional/forward model. The other one is the correct prior information of the rain.

Compared to retrieving the integral rain variables, it is more efficient to retrieve the parameters of a DSD model. The DSD is usually modeled by the gamma distribution as

$$N(D) = N_0 D^{\mu} \exp(-\Lambda D) \qquad (1),$$

where the number concentration parameter is N_0, the distribution shape parameter μ, the slope parameter Λ. Using the constraining relation by Cao et al. (2007), the gamma DSD model degrades to the two-parameter model and can be retrieved from the radar measured reflectivity at horizontal polarization (Z_H) and differential reflectivity (Z_{DR}). The constraining relation is

$$\mu = -0.0201\Lambda^2 + 0.902\Lambda - 1.718 \qquad (2).$$

The two-dimensional video disdrometer (2DVD) is capable of measuring the DSD with accuracy (Kruger and Krajewski, 2002). Since May 2005, DSD data of more than 30,000 minutes have been collected in central Oklahoma. These observations from three 2DVDs provide sufficient data to obtain valuable prior (physical and statistical) information of DSDs.

In this study, we present the results of DSD retrievals from Z_H and Z_{DR} measured by the S-band (10.7 cm) radar. The state parameters are chosen as Λ and N_0 but with appropriate forms. The prior probability density function (PDF) of state parameters and the standard deviations (SD) of Z_H and Z_{DR}, which are based on pure rain, are derived from 2DVD observations.

2. RETRIEVAL ALGORITHM

2.1 Bayesian approach

Define **x** to be the state vector, which represents the rain physics and need to be retrieved from the radar measurements. Define **y** to be the measurement vector. According to the Bayesian theorem, the posterior conditional PDF $P_{post}(\mathbf{x}|\mathbf{y})$ is given by

$$P_{post}(\mathbf{x}|\mathbf{y}) = \frac{P_f(\mathbf{y}|\mathbf{x}) \cdot P_{pr}(\mathbf{x})}{\int P_f(\mathbf{y}|\mathbf{x}) \cdot P_{pr}(\mathbf{x}) \cdot d\mathbf{x}} \qquad (3),$$

* Corresponding author address: Qing Cao; 120 David L. Boren Blvd. Suite 5900, Norman, OK 73072, U.S.A.; E-mail: qingcao@ou.edu

where $P_{pr}(\mathbf{x})$ is the prior PDF of state **x** and $P_f(\mathbf{y}|\mathbf{x})$ is the conditional PDF of the observation **y**, given the state **x**. The expected value $E(\mathbf{x})$ and the standard deviation $SD(\mathbf{x})$ are then calculated by the integration over the whole range of state **x** as

$$E(\mathbf{x}) = \frac{\int \mathbf{x} \cdot P_f(\mathbf{y}|\mathbf{x}) \cdot P_{pr}(\mathbf{x}) \cdot d\mathbf{x}}{\int P_f(\mathbf{y}|\mathbf{x}) \cdot P_{pr}(\mathbf{x}) \cdot d\mathbf{x}} \quad (4),$$

$$SD(\mathbf{x}) = \sqrt{\frac{\int (\mathbf{x} - E(\mathbf{x}))^2 \cdot P_f(\mathbf{y}|\mathbf{x}) \cdot P_{pr}(\mathbf{x}) \cdot d\mathbf{x}}{\int P_f(\mathbf{y}|\mathbf{x}) \cdot P_{pr}(\mathbf{x}) \cdot d\mathbf{x}}} \quad (5).$$

In this study, $\mathbf{x} = [N_0^*, \Lambda^*]$ and $\mathbf{y} = [Z_H, Z_{DR}]$. The N_0^* is equal to $\log_{10}N_0$ and the Λ^* is equal to $\Lambda^{0.25}$. The reason why nonlinear transformations of two DSD parameters are applied is stated in subsection 2.3.

2.2 The forward model

To simulate the radar measurements from a rain DSD, we assume the radar wavelength is 10.7 cm (S-band), the raindrop temperature is 10 °C and the standard deviation of the raindrop's canting angle equals zero. According to Zhang (et al. 2001), the Z_H, the reflectivity at vertical polarization (Z_V) and Z_{DR} are calculated as

$$Z_{H,V} = \frac{4\lambda^4}{\pi^4 |K_w|^2} \int_0^\infty |f_{a,b}(\pi)|^2 N(D)dD \quad (6),$$

and $$Z_{DR}(dB) = \log_{10}\frac{Z_H}{Z_V} \quad (7),$$

where $f_a(\pi)$ and $f_b(\pi)$ respectively represent the backscattering amplitude for horizontal and vertical polarization, λ is the wavelength, $K_w = (\varepsilon_r-1)/(\varepsilon_r+2)$ and ε_r is the complex dielectric constant of water. The $f_a(\pi)$ and $f_b(\pi)$ were calculated based on the T-matrix method and Rayleigh scattering approximation. The results of the scattering amplitude were stored as a lookup table with regard to a number of equivalent diameters. The lookup table provides a convenient way to calculate the radar variables, given a DSD.

The integral rain variables: total number concentration (N_T), median volume diameter (D_0), and rainfall rate (R) are calculated as follows.

$$N_T = \int_0^\infty N(D)dD \quad (8),$$

$$R = \frac{\pi}{6}\int_0^\infty D^3 N(D) v(D) dD \quad (9),$$

$$\int_0^{D_0} D^3 N(D)dD = \int_{D_0}^\infty D^3 N(D)dD \quad (10),$$

where v is the empirical terminal velocity proposed by Brandes (et al. 2002).

2.3 The prior distribution of DSD parameters

In order to obtain the prior information of N_0 and Λ, DSDs measured by the 2DVD were fitted to the gamma distribution using the 2nd, 4th, and 6th DSD moments, following the method of truncated moment fit described by Vivekanandan et al. (2004).

The distributions of N_0 and Λ are greatly skewed and have large dynamic ranges so that it is not appropriate to directly use N_0 and Λ as state parameters. For example, the physical property of the DSD varies greatly when Λ is small while the rain is mostly light with the narrow DSD when Λ is large (>10). That is, the physical property varies nonlinearly with Λ increasing linearly. Therefore, it is necessary to transform the evaluating range to distinguish the different physical processes. In order to reduce their dynamic ranges and mitigate the nonlinear effects, $N_0^*=\log_{10}N_0$ and $\Lambda^*=\Lambda^{0.25}$ are used. The occurrence frequency of N_0^* and Λ^* is shown in Fig 1. The dynamic ranges of N_0^* and Λ^* for light rains are reduced greatly and the dynamic ranges for moderate/heavy rains relatively account for more proportion than the ranges of N_0 and Λ. It is seen that both distributions are close to the Gaussian distributions. For this reason, N_0^* and Λ^* were chosen to be the state parameters for the retrieval.

Fig. 2 shows the contour of the occurrence frequency based on the fitted DSD parameters. The joint PDF of N_0^* and Λ^* is equal to the normalization of this distribution. It is shown that most of DSDs have the N_0 between 10^3-10^5 # m^{-3} mm^{-1} and the Λ around 2-6.

2.4 The conditional distribution

The conditional PDF $P_f(\mathbf{y}|\mathbf{x})$ is assumed to follow a bivariate-normal distribution as

$$P_f\left(Z_H, Z_{DR} \middle| \Lambda^*, N_0^*\right) =$$

$$\frac{1}{2\pi \cdot SD(Z_H) SD(Z_{DR}) \sqrt{1-\rho^2}} \exp\left(-\frac{1}{2\left(1-\rho^2\right)}\left[\frac{\left(Z_H - E(Z_H)\right)^2}{SD^2(Z_H)} - \right.\right.$$

$$\left.\left.\frac{2\rho\left(Z_H - E(Z_H)\right)\left(Z_{DR} - E(Z_{DR})\right)}{SD(Z_H) SD(Z_{DR})} + \frac{\left(Z_{DR} - E(Z_{DR})\right)^2}{SD^2(Z_{DR})}\right]\right) \quad (11),$$

where ρ is the correlation coefficient between Z_H and Z_{DR}. For rain DSD data, the correlation between Z_H and Z_{DR} is not low. As to PRD, however, Z_H and Z_{DR} errors have little correlation. As an approximation, it is reasonable to assume ρ=0 for PRD and results do not have much difference.

It is worthwhile to note that in Eq. (11) Z_H and Z_{DR} are functions of N_0^* and Λ^*, and the expected value and standard deviation are also functions of them. It is difficult to find the exact form of the latter functions. In practice, we use the DSD data to approach this problem. In order to compute the integration of Eqs. (4) and (5), we discretize N_0^* by the interval 0.1 and Λ^* by the interval 0.05. Therefore, according to the distribution shown in Fig. 2, we are able to estimate the SD and E values for each pair of discrete N_0^* and Λ^*. The contours of the estimated $SD(Z_H)$ and $SD(Z_{DR})$ are shown in Fig. 3. According to Figs. 2 and 3, the most frequent DSDs have $SD(Z_H)$ around 3 dBZ and $SD(Z_{DR})$ about 0.3 dB. We store all these values along with the prior joint PDF in the form of the lookup table for the retrieval algorithm.

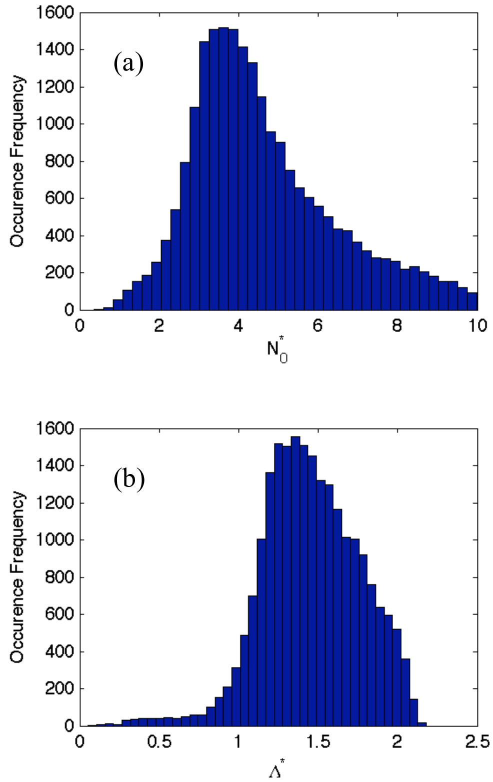

Fig. 1 Histogram of estimated DSD parameters based on 2DVD data: a) N_0^*, and b) Λ^*.

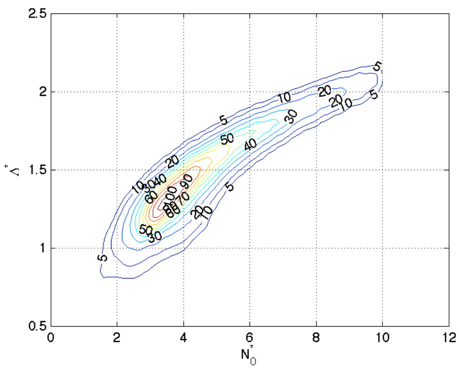

Fig. 2 Contour of the occurrence frequency of joint estimated DSD parameters.

3. RESULTS

According to DSDs measured by 2DVDs, we simulate the radar measurements Z_H and Z_{DR} and apply them to retrieve DSDs. With retrieved $E(N_0^*)$ and $E(\Lambda^*)$, mean values of rain variables are calculated. Fig. 4 shows one-one plots of retrieved values versus observations. It is seen that the values of retrieved R are very close to the observations. The table lists the bias and correlation coefficient between them, given the R< 80 mm h^{-1}. The bias of retrieved R is only -1.66% and the correlation coefficient is as high as 0.984. There are also satisfactory retrievals for variables D_0 and N_T. The scattering in Figs. 4b and 4c are mainly attributed to the constrained DSD model, which is inherently a two-parameter model only and cannot represent real DSDs perfectly. The performances of the Bayesian retrieval for R and D_0 are obviously better than the results of direct retrieval algorithm (Cao et al. 2007, Table 3). It implies that the use of prior information mitigates the error effect and improves the retrieval.

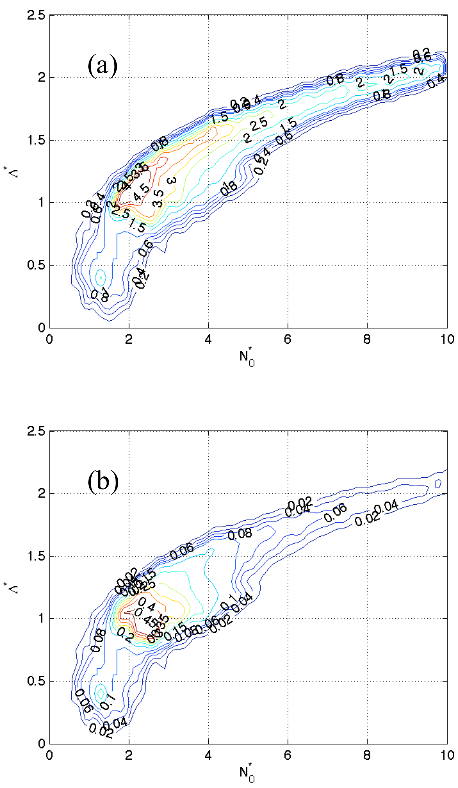

Fig. 3 Contour of the estimated standard deviations (SD) of a) Z_H (dBZ) and b) Z_{DR} (dB).

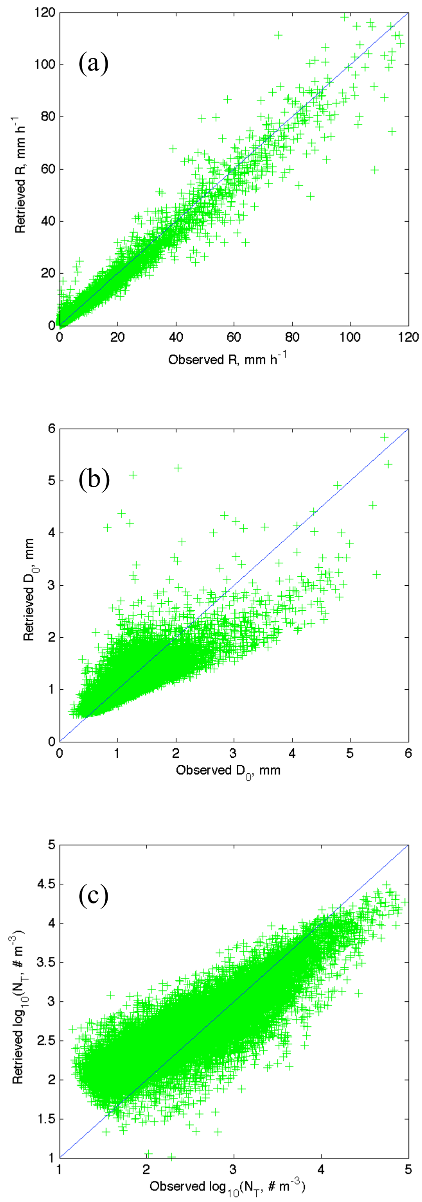

Fig. 4 One-one plot of the retrieved values versus observations: a) R, b) D_0, and c) N_T

Table: Bias and correlation coefficient for retrievals vs. observations (R< 80 mm h^{-1})

	R (mm h^{-1})	D_0 (mm)	N_T (# m^{-3})
Bias (%)	-1.66	-5.71	-19.40
Corr. Coef.	0.984	0.817	0.813

Fig. 5 shows the occurrence histogram of the estimated SD values for rain limited into different R ranges. It is seen the estimated SD value tends to decrease with the R increasing. For R<3 mm h^{-1}, the uncertainty of estimated R could be over 100%. For R>10 mm h^{-1}, SD values are normally very small, implying that the retrieval algorithm has good performance.

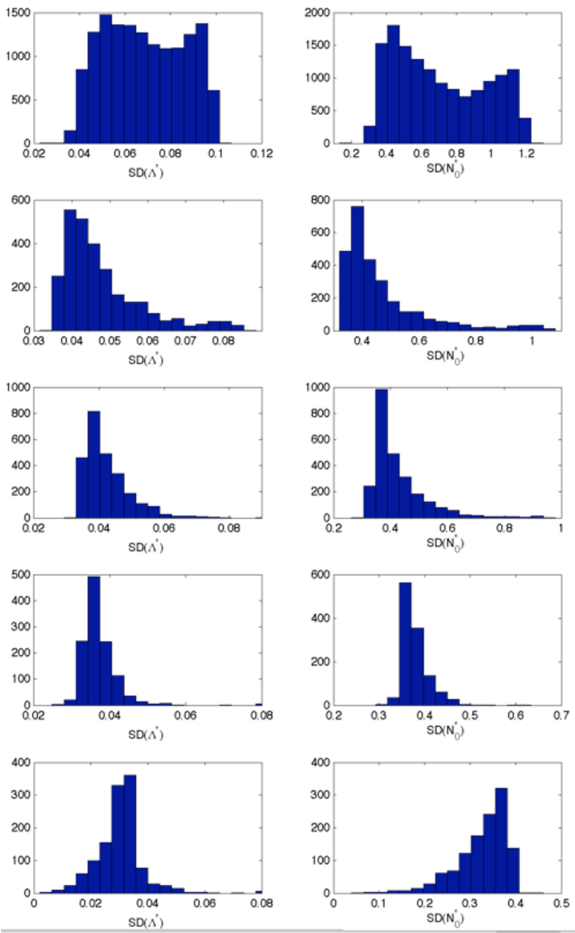

Fig. 5 Occurrence histogram of retrieved SD values. The left column is for values of SD(Λ^*) and the right column is for SD(N_0^*). The rows from up to down are for data within the range of 0<R<3, 3<R<5, 5<R<10, 10<R<20, and 20<R<100, respectively.

Fig.6 shows RHI images of a stratiform event. Z_H, Z_{DR} and the cross-correlation coefficient (ρ_{HV}) were measured by KOUN, an S-band polarimetric radar, around the 1030UTC on May 13, 2005. The hydrometeor classification is also given in Fig. 6(d). It is seen that the melting layer is located at the height of about 2-3 km. Within this region, the ρ_{HV} is mostly less than 0.9, which points out that the mixture of rain and snow/hail are present. There are also anomalous propagations within the 20 km distance on the lowest-level scanning. According to the classification, the region above melting layer is mostly the crystal and dry snow and the value of Z_{DR} is typically small, around 0.3-0.5 dB, which is much less than that of rain below the melting layer. The value of ρ_{HV}, however, is close to 1 for both regions above and below the melting layer.

Fig. 7 displays RHI images of retrieved results. As shown in Fig. 6(c) and 6(d), a clear boundary is evident at the height of 3 km. Above the height of 3 km, SD values are larger than those below the boundary. The above region includes dry snow, graupel and crystals, for which SD values are obviously larger than normal values for rain (Fig. 5). For the region of the anomalous propagation, SD values are also much larger than those for rain. These large SD values imply that the rain is not probably present there and the retrieved R and D_0 are not reliable. It is worthwhile to note that there is one exception for non-rain echoes. For the wet snow region located at the melting layer, the algorithm still gives low SD values. It is implied that other PRD, besides Z_H and Z_{DR}, are needed to discern the wet snow region.

4. SUMMARY

The Bayesian retrieval algorithm introduced in this study is an optimized algorithm. It only needs Z_H and Z_{DR} to estimate the DSD parameters and consequently other rain variables. The estimation is optimized based on the prior information provided by the 2DVD measurements. The accuracy of the 2DVD measurements and the sufficiency of the data ensure the right prior PDF. The algorithm is based on the rain model and it has the potential to distinguish the rain from other types of hydrometeors by evaluating the variance of the estimated DSD parameters.

Compared to existing algorithms of direct rain retrievals, the Bayesian algorithm gives not only the mean of DSD parameters but also their variances, which indicate the reliability of the retrieval. Compared to algorithms of variational analysis, the Bayesian algorithm has computation efficiency and is convenient to be applied for S-band radars, for which the attenuation is not significant such that the attenuation correction can be neglected.

In future work, the DSD retrieval algorithm will be improved by assimilating PRD into NWP models for the optimal retrieval using spatial and model constraints.

ACKNOWLEDGEMENT

This work was supported by a NSF grant ATM-0608168. We appreciate Dr. Schuur from NOAA's National Severe Storms Laboratory (NSSL) and Dr. Brandes from the National Center for Atmospheric Research (NCAR), who provided the part of 2DVD data. We are also thankful that the NSSL provided the KOUN radar data.

Reference:

Brandes, E. A., G. Zhang, and J. Vivekanandan, 2002: Experiments in rainfall estimation with a polarimetric radar in a subtropical environment. *J. Appl. Meteor.,* **41**, 674-685

——, ——, and ——, 2004: Drop size distribution retrieval with polarimetric radar: model and application. *J. Appl. Meteor.,* **43**, 461-475

Cao, Q., G. Zhang, E. Brandes, T. Schuur, A. Ryzhkov, and K. Ikeda, 2007: Analysis of video disdrometer and polarimetric radar data to characterize rain microphysics in Oklahoma. Accepted by *J. Appl. Meteor.*

Kruger, A., and W. F. Krajewski, 2002: Two-Dimensional Video Disdrometer: A Description. *J. Atmos. Oce. Tech.,* **19**, 602-617

Vivekanandan, J., G. Zhang, and E. Brandes, 2004: Polarimetric Radar Rain Estimators Based on Constrained Gamma Drop Size Distribution Model, *J. Appl. Meteor.*, **43**, 217-230.

Zhang, G., J. Vivekanandan, and E. Brandes, 2001: A method for estimating rain rate and drop size distribution from polarimetric radar. *IEEE Trans. Geosci. Remote Sensing,* vol. **39**, No.4, 830-840.

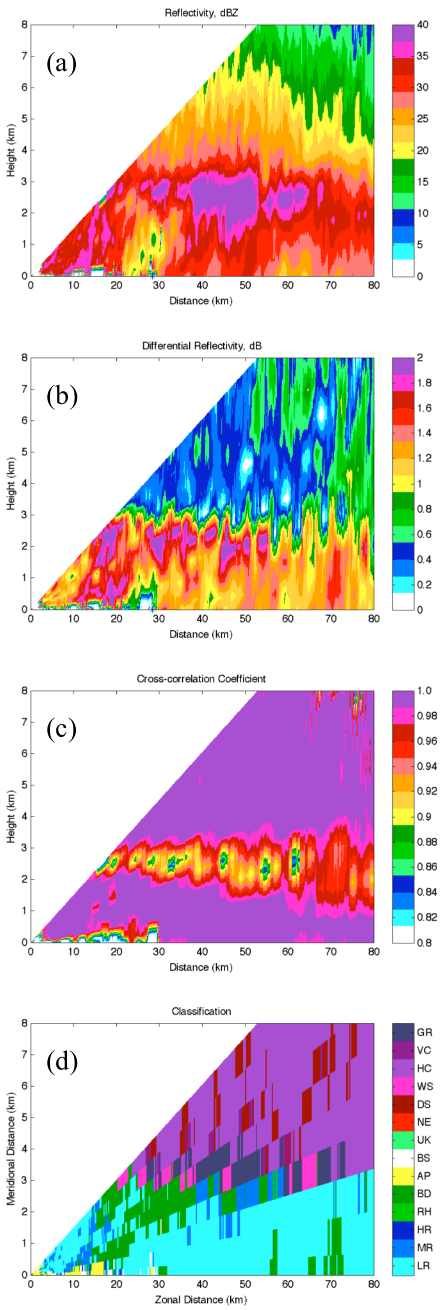

Fig. 6 RHI image of a) Z_H, b) Z_{DR},, (c) ρ_{HV} and (d) hydrometeor classification measured by KOUN radar around 1030UTC on May 13, 2005. The main classifications of hydrometeors in (d) are light rain (LR), moderate rain (MR), heavy rain (HR), rain/hail (RH), big drops (BD), dry snow (DS), wet snow (WS), horizontal crystal (HC), vertical crystal (VC), graupel (GR).

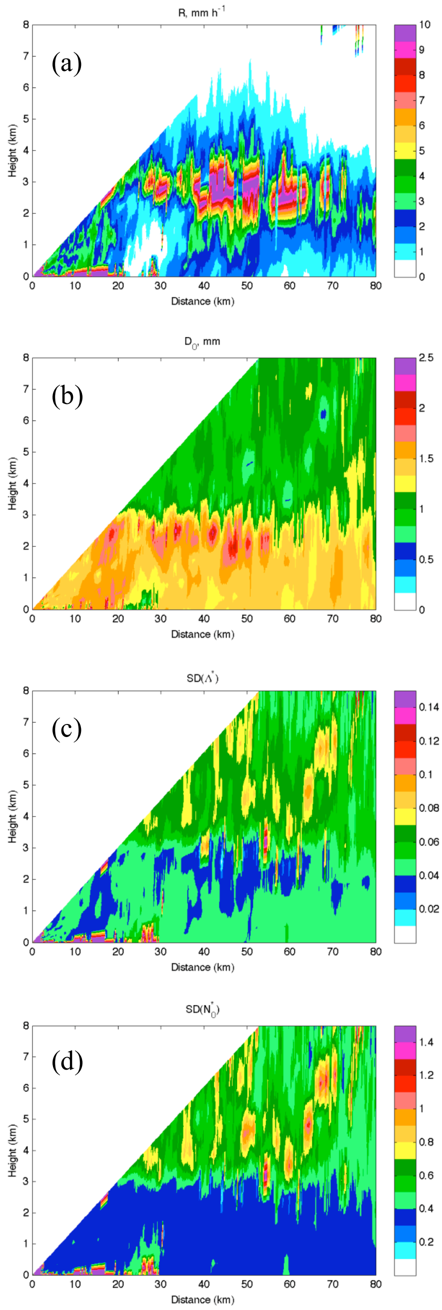

Fig. 7 RHI images of retrieved a) R (mm h^{-1}), b) D_0 (mm), c) SD(Λ^*), and d) SD(N_0^*).