JOBNAME: JTECH 00#0 2012 PAGE: 1 SESS: 8 OUTPUT: Wed Mar 14 20:58:14 2012 Total No. of Pages: 14
/ams/jtech/0/jtechD1100122

BNL-97293-2012-JA

Journal of Atmospheric and Oceanic Technology (Proof Only)

Rainfall-Rate Estimation Using Gaussian Mixture Parameter Estimator: Training and Validation

ZHENGZHENG LI AND YAN ZHANG

School of Electrical and Computer Engineering, and the Atmospheric Radar Research Center, University of Oklahoma, Norman, Oklahoma

SCOTT E. GIANGRANDE

Brookhaven National Laboratory, Upton, New York

(Manuscript received 27 July 2011, in final form 28 December 2011)

ABSTRACT

This study develops a Gaussian mixture rainfall-rate estimator (GMRE) for polarimetric radar-based rainfall-rate estimation, following a general framework based on the Gaussian mixtures model and Bayes least squares estimation for weather radar–based parameter estimations. The advantages of GMRE are 1) it is a minimum variance unbiased estimator; 2) it is a general estimator applicable to different rain regimes in different regions; and 3) it is flexible and may incorporate/exclude different polarimetric radar variables as inputs. This paper also discusses training the GMRE and the sensitivity of performance to mixture number. A large radar and surface gauge observation dataset collected in central Oklahoma during the multiyear Joint Polarization Experiment (JPOLE) field campaign is used to evaluate the GMRE approach. Results indicate that the GMRE approach can outperform existing polarimetric rainfall techniques optimized for this JPOLE dataset in terms of bias and root-mean-square error.

1. Introduction

Weather radars have been important for rainfall-rate estimation because of their ability to cover large areas with reasonable spatial and temporal resolutions. A cornerstone of radar-based rainfall estimation is the use of efficient conventional power-law relations (PLR) between the radar reflectivity factor Z and rainfall rate R [hereafter $R(Z_h)$], where Z_h is the radar reflectivity factor in horizontal polarization Z_h (mm^6 m^{-3}) and Z_H (dBZ). A common source of error for these methods is that PLR coefficients are often customized to a particular longer-term climatology or seasonal/regional precipitation regimes, and therefore are not universally applicable (e.g., Bringi et al. 2004; Cifelli et al. 2011; Fulton et al. 1998; Ryzhkov et al. 2005a; Wang and Chandrasekar 2010). Although space–time variability in the drop size distribution (DSD) contributes to such apparent diversity in power-law coefficients (Lee and Zawadzki 2005), assessing the performance of PLR estimators may be complicated further by factors that include radar attenuation in rain, beam geometrical considerations, and contamination from hail or melting layer media (e.g., Anagnostou et al. 2006; Giangrande and Ryzhkov 2008; Lee 2006).

Weather radars with dual-polarization capability provide additional insight into the precipitation medium and can help resolve some uncertainties from DSD variability and additional sources (Seliga and Bringi 1976). Dual-polarization radar moments, including the differential reflectivity (Z_{dr}) and specific differential phase (K_{dp}), allow unique insights into rain parameters, including the size, shape, and orientation of raindrops (Gorgucci et al. 2002). Algorithms that utilize polarimetric radar measurements show significant improvement over traditional $R(Z_h)$ relations and lessened sensitivity to DSD variability and partial attenuation in rain (e.g., Bringi et al. 2004; Hogan 2007; Ryzhkov et al. 2005a; Vulpiani et al. 2009). In addition to PLR methods that estimate rainfall rate R directly from polarimetric radar measurements, recent studies have approached the

Corresponding author address: Yan Zhang, 110 W. Boyd Street, Devon Energy Hall 150, University of Oklahoma, Norman, Norman, OK 73019.
E-mail: rockee@ou.edu

DOI: 10.1175/JTECH-D-11-00122.1

© 2012 American Meteorological Society

JOBNAME: JTECH 00#0 2012 PAGE: 2 SESS: 8 OUTPUT: Wed Mar 14 20:58:14 2012 Total No. of Pages: 14
/ams/jtech/0/jtechD1100122

Journal of Atmospheric and Oceanic Technology (Proof Only)

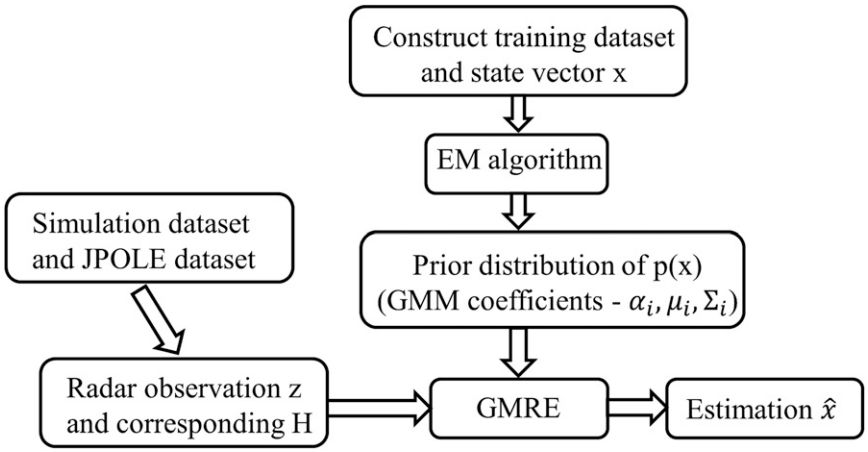

FIG. 1. Flowchart of Gaussian mixture rainfall-rate estimator.

issue of DSD variability by computing the corresponding R from DSD parameters retrieved from polarimetric radar measurements (e.g., Vulpiani et al. 2006; Cao et al. 2010) using neural network or Bayesian approaches.

Most of the polarimetric rainfall-rate estimators are still deterministic PLRs, where Z_h, Z_{dr} and K_{dp} are used in different combinations or the most appropriate estimator is selected for a given set of polarimetric radar measurements (Ryzhkov et al. 2005a). Deterministic estimators usually fail to account for the fact that microphysics varies in space and time, even within the same precipitation, leading to estimates that are less than optimal. To address this problem, Bringi et al. (2004) derived a new $R(Z_h)$ that varies continuously in space and time. Hogan (2007) presented a spatially variational method where coefficient a in $R(Z_h)$ is iteratively refined. Vulpiani et al. (2009) developed a nonlinear estimator based on neural network, and Cao et al. (2010) proposed a Bayesian approach where only Z_h and Z_{dr} are used due to the assumption of a single Gaussian distribution of the joint distribution of Z_h and Z_{dr} given DSD parameters.

Recently, Li and Zhang (2011) introduced the Gaussian mixture parameter estimator (GMPE), a linear Bayesian estimator. Because the Gaussian mixture model (GMM) ensures convergence to the prior distribution of dual-polarization variables, the GMPE (a minimum variance unbiased estimator) was shown to outperform PLRs in both rainfall-rate estimation and attenuation correction using simulated polarimetric radar measurements. However, the performance of the GMPE has not been tested for real-world rainfall applications.

Rainfall-rate estimators can be developed either through measurements or simulations. Rainfall-rate estimators developed from measurements are usually optimized for specific radar and regions but are less suitable for others, because precipitation is different for different regions, and radars may have different calibration errors and noise levels. On the other hand, rainfall-rate estimators developed from simulations are more general and less sensitive to measurement error, but they depend on different assumptions. Because of the natural variability of raindrop size, shape, and terminal velocity, any specific model–assumption might lead to errors in simulated polarimetric radar measurements. The single-cell Monte Carlo simulation proposed in Li et al. (2011) and adopted in this study addresses such variability by allowing variables, including raindrop shape, canting angle, and DSD, to have uncertainties, making them more suitable for GMPE development and enabling them to better embody the simulated dual-polarization variable distribution.

In this study, the GMPE approach is applied to polarimetric radar-based rainfall-rate estimation. To distinguish it from general GMPE, it is renamed the Gaussian mixture rainfall-rate estimator (GMRE). The flowchart of the GMRE approach is shown in Fig. 1. The GMRE approach was validated by using data collected during the Joint Polarization Experiment (JPOLE) from the well-gauged central Oklahoma region and S-band radar data from the KOUN radar (Doviak et al. 2002), the polarimetric prototype of the Weather Surveillance Radar-1988 Doppler (WSR-88D), over a multiyear period (Ryzhkov et al. 2005b). Performance of GMRE will be compared to other rainfall-rate estimators that were developed and tested on the JPOLE dataset.

This paper is organized as follows: Sections 2 and 3 introduce the general background and details of the Monte Carlo simulation as well as the GMRE algorithm. Section 4 gives a description of the JPOLE dataset and processing methods, followed by the results and comparison of GMRE and other rainfall-rate estimators in section 5. Section 6 presents the discussion and conclusions.

JOBNAME: JTECH 00#0 2012 PAGE: 3 SESS: 8 OUTPUT: Wed Mar 14 20:58:17 2012 Total No. of Pages: 14
/ams/jtech/0/jtechD1100122

Journal of Atmospheric and Oceanic Technology (Proof Only)

2. Monte Carlo simulation

a. Introduction

Polarimetric variables and the associated rain parameters for a given radar resolution volume are influenced by several factors, including DSD, drop shape behavior, drop canting angles, and terminal velocity of a raindrop. Unlike simulations that employ fixed relations, Monte Carlo simulation allows variables to have uncertainties (or randomness) to avoid assumptions or loss of generality. In this way, Monte Carlo simulation outputs can capture the desirable variable distributions beyond an ensemble average value and include important statistical information.

For this study, a single-cell Monte Carlo simulation can be considered as a single volume (cell) filled with uniformly distributed raindrops. Outputs from the simulation are the composite of echoes backscattered by raindrops. The number and size of raindrops in the volume are controlled by the DSD, which is usually represented by models that contain free parameters. This study adopts an exponential distribution as given by

$$N(D) = N_0 \exp(-\Lambda D). \tag{1}$$

This distribution depends on the following two parameters: the intercept parameter N_0 ($\mathrm{m^{-3}\ mm^{-1}}$) and slope parameter Λ ($\mathrm{mm^{-1}}$). Slope is uniquely determined if N_0 and water content W are known when given by

$$\Lambda = \left(\frac{\pi N_0 \rho_w}{W}\right), \tag{2}$$

where ρ_w is the density of water. Because N_0 and W have physical meaning, the dynamic range of both is well studied, and only a weak correlation is found between those parameters (Zhang et al. 2008). Even though the exponential distribution may not represent very small or large raindrops as well as the Gamma distribution (three free parameters), selecting this distribution helps reduce the number of unrealistic parameter cases. Once a DSD is obtained, rainfall rate R ($\mathrm{mm\ h^{-1}}$) can be computed from

$$R = 6\pi \times 10^{-4} \int_{D_{\min}}^{D_{\max}} v(D) D^3 N(D)\, dD, \tag{3}$$

where $v(D)$ is the terminal velocity relationship from Brandes et al. (2002).

To further reduce the DSD error on small raindrops and include enough big drops, $D_{\min}$ and $D_{\max}$ are set at 0.5 and 8 mm, respectively. Raindrops are modeled as oblate spheroids with the polynomial relation between axis ratio r_a (minor to major axis) and equivolume diameter D, as given in Brandes et al. (2002). To further generalize for different kinds of raindrop behavior, randomness between $[-0.2(1 - r_a), 0.2(1 - r_a)]$ is added to r_a (Li et al. 2011). Scattering amplitudes of a raindrop are calculated using the T-matrix method (Mishchenko 2000). If the copolar backscattering and forward-scattering amplitudes are denoted as $S^b_{\mathrm{hh,vv}}$ and $S^f_{\mathrm{hh,vv}}$, then the copolar radar reflectivity factor of horizontal (Z_h) and vertical polarization (Z_v), differential reflectivity (Z_{dr}), and specific differential phase (K_{dp}) of the volume are defined as follows:

$$Z_{h,v} = \frac{4\lambda^2}{\pi^4 |K_w|^2 V} \sum S^b_{\mathrm{hh,vv}}, \tag{4}$$

$$Z_{\mathrm{dr}} = \frac{Z_h}{Z_v}, \tag{5}$$

$$K_{\mathrm{dp}} = 10^{-3} \frac{180}{\pi V} \sum \Re(S^f_{\mathrm{hh}} - S^f_{\mathrm{vv}}), \tag{6}$$

where V is the size of the volume and $\sum$ indicates summation over all raindrops in the volume. To ensure that enough raindrops are in the volume and balance computational load, V is set at 1000 $\mathrm{m^3}$.

b. Simulation procedures

Shown by many observation studies, intercept parameter and water content vary for different rain regimes (e.g., Waldvogel 1974; Zhang et al. 2008). Reciprocally, different types of rain may be emulated from randomly generating N_0 and W. The empirical range of N_0 is from $10^{1.5}$ to 10^6 $\mathrm{m^{-3}\ mm^{-1}}$, while W can reach 10 $\mathrm{g\ m^{-3}}$ (Zhang et al. 2008). While the ranges of N_0 and W are well studied, the distribution of N_0 and W remains less certain. In some studies (e.g., Li et al. 2011; Vulpiani et al. 2006), a uniform distribution of DSD parameters is assumed, which leads to equal probability for different rain types. This assumption may not hold in general because smaller rainfall ($R < 30{-}40$ $\mathrm{mm\ h^{-1}}$) is more frequent than heavier rainfall intervals/cases.

In this paper, prior distributions of N_0 and W are designed to favor rainfall lower than 40 $\mathrm{mm\ h^{-1}}$ and marginalize the probability of extreme rain cases by setting W from a one-sided Gaussian distribution and N_0 from a uniform distribution with a smaller upper bound. Table 1 gives details of the simulation. Outputs of the T1 simulation include rain microphysics parameters N_0, Λ, R, and the corresponding dual-polarization variables Z_h, Z_{dr}, and K_{dp}. It is worth mentioning that 8000 cases have been generated to help provide statistical significance. As illustrated in Fig. 2, in the majority of the cases R is F2

JOBNAME: JTECH 00#0 2012 PAGE: 4 SESS: 8 OUTPUT: Wed Mar 14 20:58:17 2012 Total No. of Pages: 14
/ams/jtech/0/jtechD1100122

Journal of Atmospheric and Oceanic Technology (Proof Only)

TABLE 1. Key parameters of the single-cell Monte Carlo simulation.

Radar frequency	2.705 GHz (matching KOUN frequency)
Radar elevation angle	0°
Volume of a single radar resolution cell	1000 m^3
Temperature	Uniformly distributed between 5° and 20°C
Raindrop size	Between 0.5 and 8 mm with a step size of 0.1 mm
Canting angle distribution	Gaussian with a 0 mean and standard deviation 10°
Scattering model	T matrix
Axis ratio of raindrops	(Brandes et al. 2002)
Terminal velocity of raindrops	(Brandes et al. 2002)
Intercept parameter N_0 (m^{-3} mm^{-1})	Uniformly distributed between $10^{1.5}$ and 10^4
Water content W (g mm^{-3})	One-sided Gaussian with mean 0 and standard deviation 2

lower than 40 mm h^{-1} and the number of occurrences decreases significantly as R increases. Even though the prior distribution input into the Monte Carlo simulation emphasizes smaller rainfall, a broad range of rainfall is still covered because R reaches as high as 180 mm h^{-1}. F3 Figures 3a,b show the scatterplots of Z_H and Z_{DR} as well as Z_H and K_{dp} from the MC simulation. According to the Next Generation Weather Radar (NEXRAD) the $R(Z_h)$ relationship [Eq. (18)], $R = 20$ mm h^{-1} corresponds to approximately 43 dBZ for observed reflectivity. Combined with the R distribution in Fig. 2, where 70% of the occurrences are $R > 20$ mm h^{-1}, the majority of cases concentrated between 43 and 60 dBZ can be explained. Because of a large amount of big and oblate raindrops caused by a combination of large R and N_0 as well as randomness added to the axis ratio relation, there are some extreme cases where $Z_H >$ 55 dBZ, $Z_{DR} > 5$ dB, and $K_{dp} > 2.5°$ km^{-1} in the simulation dataset. An advantage of the Monte Carlo simulation is that it can provide the relative possibility of occurrence for extreme cases. The incorporation of extreme cases are necessary for training the GMM and it will not influence the performance of GMRE, because the GMM always converges to the true distribution as the number of mixtures increases [GMM and number of F4 mixtures is defined in Eq. (7)]. Figures 4a,b present the approximate distribution of Z_H and Z_{DR} from one trained GMM with 5 mixtures and another one with 20 mixtures. The approximate distribution from the GMM with 20 mixtures clearly shows more detail and is much closer to the original distribution in the simulation dataset (Fig. 3a). In contrast, the approximate distribution from the GMM with five mixtures ignores some details while preserving the key portions of the original distribution.

3. Gaussian mixture rainfall-rate estimator

If polarimetric radar measurements are denoted as vector $\mathbf{z}$, rainfall-rate retrieval is based on the connections between $\mathbf{z}$ and R. There are mainly three kinds of rainfall-rate retrieval approaches. Many conventional and polarimetric approaches, including NEXRAD, assume PLRs between $\mathbf{z}$ and R and use linear regression models (e.g., Bringi et al. 2004; Cifelli et al. 2011; Ryzhkov et al. 2005a). Neural network approaches consider a black box that has $\mathbf{z}$ as input and R as output (e.g., Vulpiani et al. 2006, 2009). Bayesian probability approaches try to estimate R from maximizing the posterior probability $p(R \mid \mathbf{z})$ (e.g., Cao et al. 2010; Chiu and Petty 2006; Di Michele et al. 2005; Evans et al. 1995). This section presents the theoretical fundamentals of GMRE as well as training and testing GMRE using simulation dataset, following Li and Zhang (2011).

a. *Theoretical fundamentals of the GMRE approach*

Microphysics parameters such as R, W, N_0, and Λ and the corresponding radar variables such as Z_h, Z_{dr}, and K_{dp} of a radar resolution volume (single cell) can be combined and considered as an unknown and random vector (called state vector) $\mathbf{x}$, such as $\mathbf{x} = (R, Z_h, Z_{dr})$.

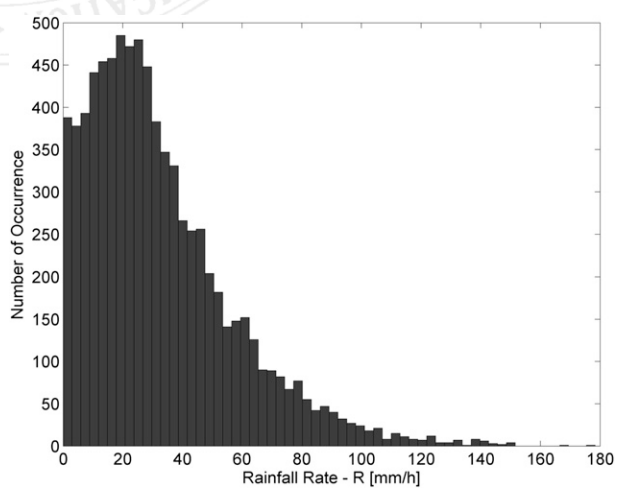

FIG. 2. Distribution of R in the Monte Carlo simulation.

JOBNAME: JTECH 00#0 2012 PAGE: 5 SESS: 8 OUTPUT: Wed Mar 14 20:58:18 2012 Total No. of Pages: 14
/ams/jtech/0/jtechD1100122

Journal of Atmospheric and Oceanic Technology (Proof Only)

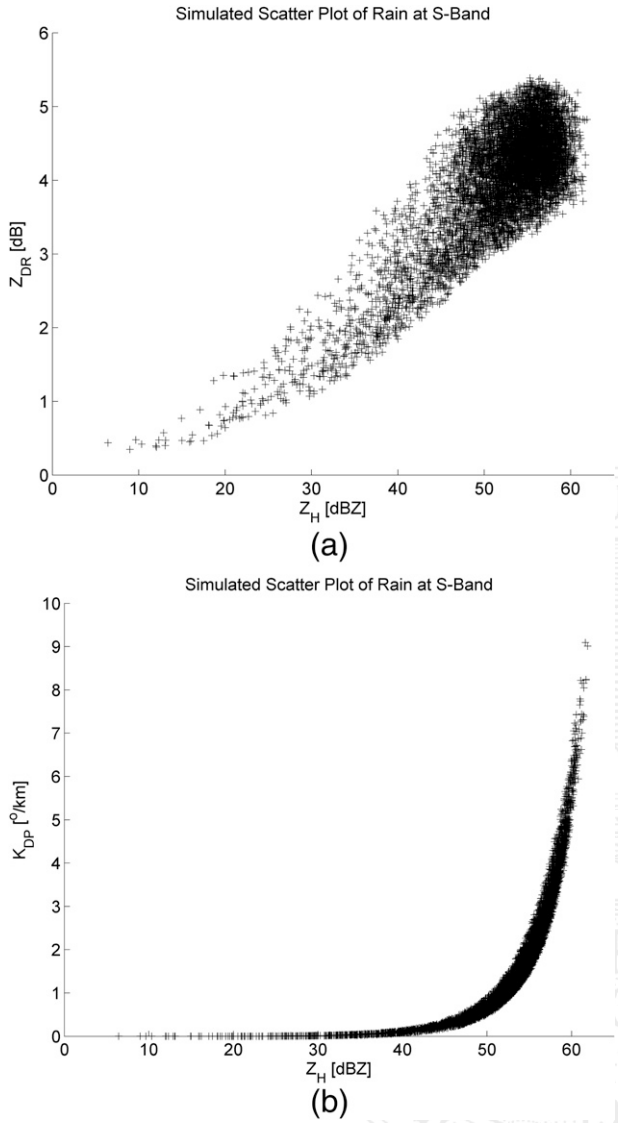

FIG. 3. Scatterplots of S-band dual-polarized radar signatures from Monte Carlo simulation: (a) Z_{DR} and (b) K_{dp}.

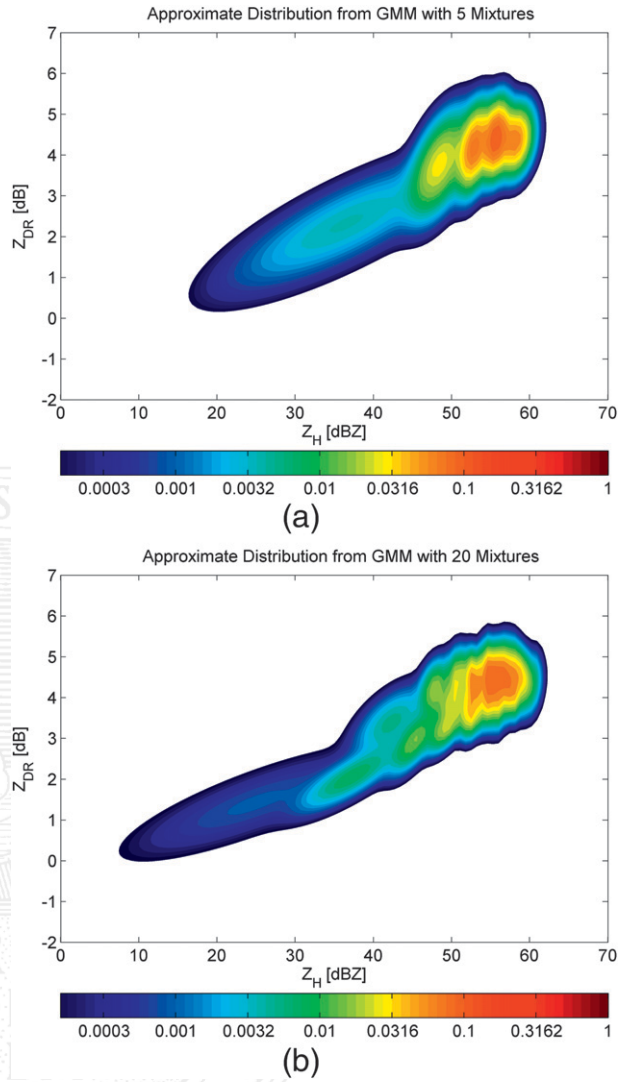

FIG. 4. Approximate distribution from GMM for simulated radar signatures: approximate distribution from GMM with (a) 5 and (b) 20 mixtures. In both plots, warmer color represents higher probability density.

The prior distribution of $\mathbf{x}$, $p(\mathbf{x})$, can be learned and represented by the GMM and is expressed as

$$p(\mathbf{x}) = \sum_{i=1}^{M} \alpha_i \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), \tag{7}$$

where $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$; M is the number of Gaussian mixtures used; and α_i, $\boldsymbol{\mu}_i$, and $\boldsymbol{\Sigma}_i$ are the weighting, mean, and covariance matrix for the ith Gaussian mixture, respectively. With a given number of mixtures, the GMM can be trained using the expectation maximization (EM) algorithm (Russell and Norvig 2009) from training datasets.

Given radar measurements $\mathbf{z}$ and measurement noise vector $\mathbf{v}$, the estimation problem can be formulated based on a linear relationship,

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{v}. \tag{8}$$

Matrix $\mathbf{H}$ links the state vector to observation and can be easily modified to include/exclude different sensor output variables. For example, if a GMM is trained with state vector $\mathbf{x} = (R, Z_h, Z_{dr}, K_{dp})$, then GMRE can be used with different $\mathbf{H}$ regardless of whether K_{dp} is available [$\mathbf{z} = (Z_h, Z_{dr}, K_{dp})$ or not $\mathbf{z} = (Z_h, Z_{dr})$], and remains the best estimator for both scenarios.

JOBNAME: JTECH 00#0 2012 PAGE: 6 SESS: 8 OUTPUT: Wed Mar 14 20:58:32 2012 Total No. of Pages: 14
/ams/jtech/0/jtechD1100122

Journal of Atmospheric and Oceanic Technology (Proof Only)

According to the Bayesian theorem, conditional distribution $p(\mathbf{x} \mid \mathbf{z})$, also known as the posterior distribution, yields

$$p(\mathbf{x}\,|\,\mathbf{z}) = \frac{p(\mathbf{z}\,|\,\mathbf{x})p(\mathbf{x})}{p(\mathbf{z})}. \tag{9}$$

If measurement noise $\mathbf{v}$ is modeled as white Gaussian noise from $\mathcal{N}(0, \mathbf{R})$, $p(\mathbf{z} \mid \mathbf{x})$ is

$$p(\mathbf{z}\,|\,\mathbf{x}) = \mathcal{N}(\mathbf{z} - \mathbf{H}\mathbf{x}; 0, \mathbf{R}) = \mathcal{N}(\mathbf{z}; \mathbf{H}\mathbf{x}, \mathbf{R}) \tag{10}$$

and $p(\mathbf{z})$, according to linear transformation and the addition property of Gaussian distribution, is also a Gaussian mixture with the same number of mixtures as $p(\mathbf{x})$ and can be written as

$$p(\mathbf{z}) = \sum_{i=1}^{M} \alpha_i \mathcal{N}(\mathbf{z}; \mathbf{H}\boldsymbol{\mu}_i, \mathbf{P}_i), \tag{11}$$

where $\mathbf{P}_i = \mathbf{H}\boldsymbol{\Sigma}_i\mathbf{H}^{\mathrm{T}} + \mathbf{R}$ is the covariance matrix for ith Gaussian mixture in $p(\mathbf{z})$. Plugging $p(\mathbf{z} \mid \mathbf{x})$, $p(\mathbf{x})$ and $p(\mathbf{z})$ into Eq. (9) yields

$$p(\mathbf{x}\,|\,\mathbf{z}) = \frac{\mathcal{N}(\mathbf{z}; \mathbf{H}\mathbf{x}, \mathbf{R}) \sum_{i=1}^{M} \alpha_i \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)}{\sum_{i=1}^{M} \alpha_i \mathcal{N}(\mathbf{z}; \mathbf{H}\boldsymbol{\mu}_i, \mathbf{P}_i)}. \tag{12}$$

Because the product of two Gaussian distributions is still Gaussian (see the appendix),

$$\mathcal{N}(\mathbf{z}; \mathbf{H}\mathbf{x}, \mathbf{R})\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = \mathcal{N}(\mathbf{z}; \mathbf{H}\boldsymbol{\mu}_i, \mathbf{P}_i)\mathcal{N}(\mathbf{x}; \hat{\boldsymbol{\mu}}_i, \hat{\mathbf{P}}_i), \tag{13}$$

where $\hat{\boldsymbol{\mu}}_i = \boldsymbol{\mu}_i + \mathbf{K}_i(\mathbf{z} - \mathbf{H}\boldsymbol{\mu}_i)$, $\hat{\mathbf{P}}_i = (\mathbf{I} - \mathbf{K}_i\mathbf{H})\boldsymbol{\Sigma}_i$, and $\mathbf{K}_i = \boldsymbol{\Sigma}_i\mathbf{H}^{\mathrm{T}}\mathbf{P}_i^{-1}$. If β_i is set to be

$$\beta_i = \frac{\alpha_i \mathcal{N}(\mathbf{z}; \mathbf{H}\boldsymbol{\mu}_i, \mathbf{P}_i)}{\sum_{i=1}^{M} \alpha_i \mathcal{N}(\mathbf{z}; \mathbf{H}\boldsymbol{\mu}_i, \mathbf{P}_i)}, \tag{14}$$

then Eq. (12) can be further written as

$$p(\mathbf{x}\,|\,\mathbf{z}) = \sum_{i=1}^{M} \beta_i \mathcal{N}(\mathbf{x}; \hat{\boldsymbol{\mu}}_i, \hat{\mathbf{P}}_i). \tag{15}$$

As shown in Eq. (15), $p(\mathbf{x} \mid \mathbf{z})$ is also a Gaussian mixture with the same number of mixtures as $p(\mathbf{x})$ and β_i is the weighing of the ith Gaussian mixture in $p(\mathbf{x} \mid \mathbf{z})$. The Bayes least squares estimate of $\mathbf{x}$ is given as the conditional mean (Lewis et al. 2006)

$$\hat{\mathbf{x}} = E[\mathbf{x}\,|\,\mathbf{z}] = \sum_{i=1}^{M} \beta_i \hat{\boldsymbol{\mu}}_i, \tag{16}$$

where $E[\cdot]$ is the expectation operator. Here, $\hat{\mathbf{x}}$ is considered the best estimate of $\mathbf{x}$ in terms of minimum variance and unbiased performance. Estimates of R, N_0, and Λ are obtained at the same time using the same GMRE model.

b. Training of GMREs

Because rainfall rate R can be estimated directly from radar observations or recovered from DSD parameters N_0 and Λ, the state vector is set as $\mathbf{x} = (R, N_0, \Lambda, Z_H, Z_{\mathrm{DR}}, K_{\mathrm{dp}})^{\mathrm{T}}$ to compare the performance of both approaches. Even though other dual-polarization variables, such as the linear depolarization ratio (LDR) and correlation coefficients (ρ_{hv}), are not included in the state vector in this study, GMRE can discover and use hidden relationships among different variables, and additional variables would generally lead to a better performance.

Because a GMM would converge to any particular distribution if sufficient mixtures are used, it is safe to assume that prior distribution $p(\mathbf{x})$ has a Gaussian mixture form. Training GMRE is a learning process during which knowledge of $p(\mathbf{x})$ is acquired from the training dataset, and it is stored in the weightings, means, and covariance matrixes of a GMM. Therefore, the training problem becomes how to estimate some unknown parameters (α_i, $\boldsymbol{\mu}_i$, and $\boldsymbol{\Sigma}_i$) from a given dataset. The EM algorithm (Russell and Norvig 2009), an iterative optimization method, has been widely used in mixture models. It proceeds as follows: in the E step, a posterior probability is assigned to each individual sample in the dataset. In the M step, a new estimate of the unknown parameters is obtained by increasing the global likelihood, that is, the combination–product of all posterior probabilities. The algorithm will continue until the global likelihood can no longer be increased and it converges to the nearest local maximum, which depends on the initial clustering values of the dataset. Because of the limited number of mixtures and the convergence to the local maximum of the EM algorithm in training GMM, while GMRE is the optimized estimator in theory, training GMRE [or, in other words, characterizing $p(\mathbf{x})$] may lead to a suboptimum performance of GMRE. However, near-optimal performance can be achieved by choosing the proper number of mixtures and better initial clustering values. In this study, the k-means clustering algorithm (Russell and Norvig 2009) is used in the initial clustering of the training datasets for better initial clustering values and faster convergence in GMM training.

c. Results of GMREs on simulation datasets

Simulation datasets are divided into two portions: 7000 cases are used for training and the remaining 1000 cases are used for testing GMRE. To better analyze the

JOBNAME: JTECH 00#0 2012 PAGE: 7 SESS: 8 OUTPUT: Wed Mar 14 20:58:32 2012 Total No. of Pages: 14
/ams/jtech/0/jtechD1100122

Journal of Atmospheric and Oceanic Technology (Proof Only)

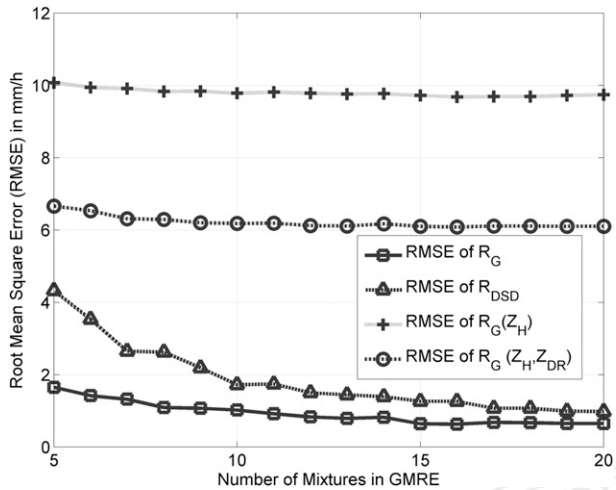

FIG. 5. RMSEs of GMRE with different inputs as the number of mixtures increases.

performance and compare it with other approaches, three error terms, including bias, RMSE, and fractional standard error (FSE) are considered. If X is the parameter (such as R, N_0, and Λ) being estimated, then $\hat{X}$ is the estimated value of the parameter and the estimation error is defined as $\epsilon_X = \hat{X} - X$. Thus, the bias is the mean of estimation error $\bar{\epsilon}_X$, RMSE can be calculated from $\sqrt{\bar{\epsilon}_X^2 + \sigma_X^2}$, and FSE (%) yields

$$\mathrm{FSE} = 100\frac{\mathrm{RMSE}}{\overline{X}}, \tag{17}$$

where σ_X is the standard deviation of the estimation and $\overline{X}$ is mean of the parameter.

Once the GMRE has been constructed, it is ready to be tuned and perform in different scenarios. For example, if only reflectivity factor Z_H is available, such as with the legacy WSR-88D, GMRE can be used with $\mathbf{H} = [0, 0, 0, 1, 0, 0]$. The rainfall rate retrieved by GMRE with Z_H will be denoted as $R_G(Z_H)$. For dual-polarized radar without (or with low quality) differential phase measurements, input to GMRE becomes $\mathbf{z} = (Z_H, Z_{\mathrm{DR}})^{\mathrm{T}}$ and $\mathbf{H}$ yields

$$\mathbf{H} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

The rainfall rate retrieved with $\mathbf{z} = (Z_H, Z_{\mathrm{DR}})^{\mathrm{T}}$ will be denoted as $R_G(Z_H, Z_{\mathrm{DR}})$. For radars with full dual-polarization capabilities ($\mathbf{z} = (Z_H, Z_{\mathrm{DR}}, K_{\mathrm{dp}})^{\mathrm{T}}$), the same GMRE also can be applied. With $\mathbf{z} = (Z_H, Z_{\mathrm{DR}}, K_{\mathrm{dp}})^{\mathrm{T}}$, R can be directly estimated from GMRE (denoted as R_G) or calculated from retrieved DSD parameters (N_0 and Λ) using Eq. (3) (denoted as R_{DSD}).

TABLE 2. Rain parameters retrieved by GMREs with 5 mixtures for the simulation dataset: N_0 (mm^{-1} m^{-3}), Λ (mm^{-1}), and all rainfall rate R (mm h^{-1}).

	$\log_{10}(N_0)$	Λ	$R_G(Z_H)$	$R_G(Z_H, Z_{\mathrm{DR}})$	R_{DSD}	R_G
Bias	0.17	0.01	−0.16	−0.27	0.2	0.07
STD	1.47	0.11	10.07	6.66	4.31	1.65
RMSE	1.48	0.11	10.07	6.66	4.32	1.65
FSE (%)	2.96	12.96	30.93	20.47	13.26	5.08

Figure 5 illustrates the RMSEs of GMRE with different inputs and number of mixtures. In general, more observation variables input into GMRE would lead to better performance. As the number of mixtures increases, the RMSEs of $R_G(Z_H)$, $R_G(Z_H, Z_{\mathrm{DR}})$, and R_G improve slowly while the RMSE of R_{DSD} significantly lowers from more than 4 mm h^{-1} to less than 2 mm h^{-1}. Tables 2 and 3 compare the performance of GMRE with 5 and 20 mixtures. The GMRE with 20 mixtures is better than the GMRE with 5 mixtures in basically every category. As mentioned in last section, GMRE is a minimum variance, unbiased estimator as long as GMM converged to prior distribution $p(\mathbf{x})$. More mixtures in GMM leads to a closer approximate distribution to $p(\mathbf{x})$ (as can be seen in Figs. 4a,b) and better estimation performance, which would eventually reach minimum variance and unbiased estimations. Therefore, the question becomes how many mixtures are appropriate for GMRE, and the answer varies for different applications. For $R_G(Z_H)$ and $R_G(Z_H, Z_{\mathrm{DR}})$, GMRE with 5 mixtures would be sufficient to perform near its optimal point (with minimum variance and unbiased estimation), while R_G needs 15 mixtures and R_{DSD} may need more than 20 to reach to their optimal performance on the simulation dataset. Figure 6 illustrates the plots of the rainfall-rate estimation from R_G and R_{DSD} with 5 and 20 mixtures versus the simulated truth data. Given the same weather radar observations $\mathbf{z} = (Z_H, Z_{\mathrm{DR}}, K_{\mathrm{dp}})^{\mathrm{T}}$, R_G performs significantly better than R_{DSD} when GMRE has five mixtures because R_{DSD} is calculated from the retrieved N_0 and Λ, where the estimation error of N_0 and Λ accumulates and magnifies, therefore leading to larger RMSE for R_{DSD}. For GMRE with 20 mixtures, the better performance of R_{DSD} is obtained

TABLE 3. Rain parameters retrieved by GMREs with 20 mixtures for the simulation dataset: N_0 (mm^{-1} m^{-3}), Λ (mm^{-1}), and all rainfall rate R (mm h^{-1}).

	$\log_{10}(N_0)$	Λ	$R_G(Z_H)$	$R_G(Z_H, Z_{\mathrm{DR}})$	R_{DSD}	R_G
Bias	0.08	0	−0.05	−0.01	0.12	0.03
STD	0.68	0.07	9.74	6.1	0.98	0.65
RMSE	0.69	0.07	9.74	6.1	0.98	0.65
FSE (%)	1.37	8.31	29.9	18.75	3.02	2.01

JOBNAME: JTECH 00#0 2012 PAGE: 8 SESS: 8 OUTPUT: Wed Mar 14 20:58:33 2012 Total No. of Pages: 14
/ams/jtech/0/jtechD1100122

Journal of Atmospheric and Oceanic Technology (Proof Only)

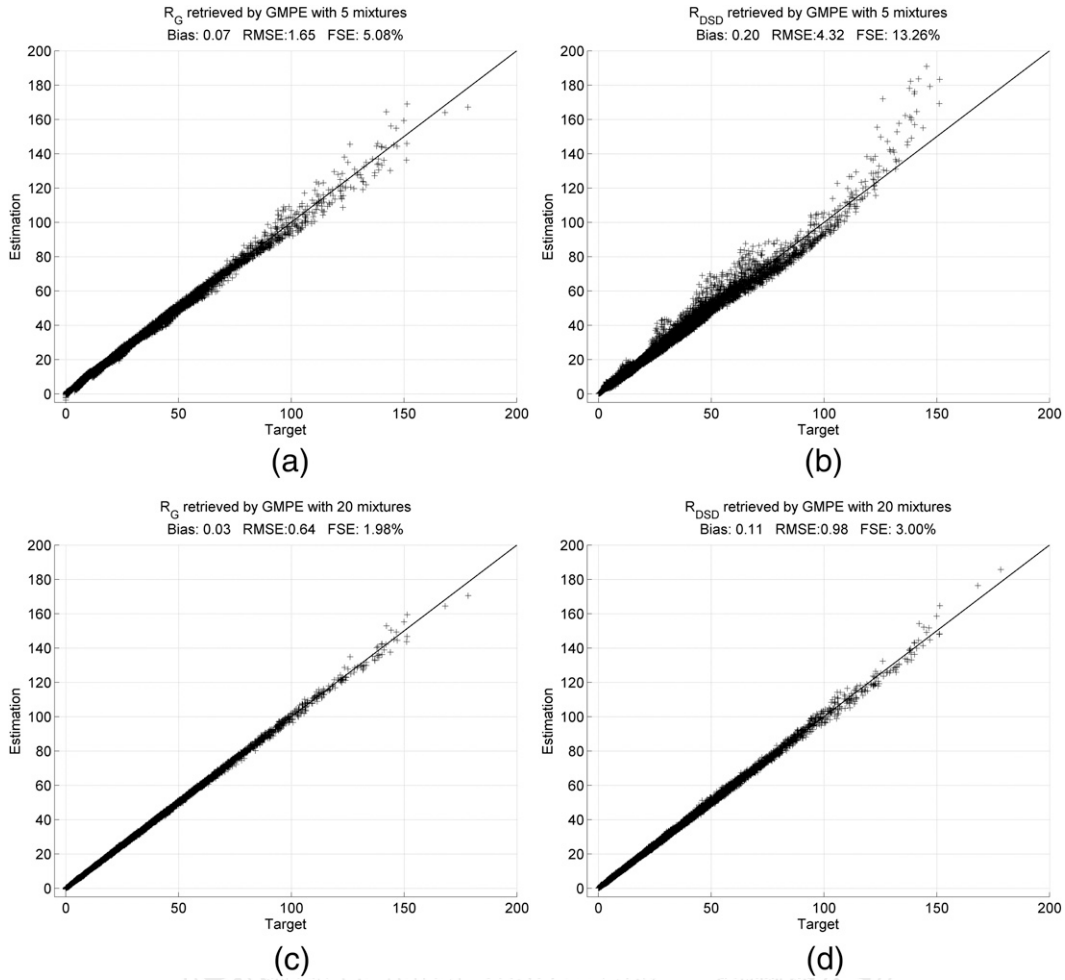

FIG. 6. Plots of estimated rainfall rate and true rainfall rate for simulation dataset: (a) R_G retrieved by GMRE with 5 mixtures, (b) R_{DSD} retrieved by GMRE with 5 mixtures, (c) R_G retrieved by GMRE with 20 mixtures, and (d) R_{DSD} retrieved by GMRE with 20 mixtures. Bias and RMSE are in mm h^{-1} for all plots.

due to more accurate estimates of N_0 and Λ. As the number of mixtures increases, the performance of R_{DSD} improves, but it will not surpass $R_{G.}$ Therefore, R_{DSD} will not be considered in later discussion.

The above simulation results indicate that if GMRE is trained from a dataset whose distribution matches that of the testing dataset, then the GMRE with more mixtures has better performance because GMM converges closer to the true distribution as the number of mixtures increases. Also, these simulation results assume a noise-free environment, which means the noise covariance matrix $\mathbf{R} = 0$. In terms of the performance of the GMRE in a noisy environment, it has been shown that estimation from GMRE is robust when the observations contain Gaussian noise (Li and Zhang 2011). As long as the Gaussian noise assumption holds, then GMRE would remain a minimum variance and unbiased estimator.

4. JPOLE dataset description

The JPOLE dataset is a polarimetric radar dataset collected between 2002 and 2005 in central Oklahoma using the KOUN WSR-88D quality radar. A total of 43 events of various precipitation types, including warm-season convective storms containing hail, mesoscale convective systems (MCS) with intense squall lines and trailing stratiform precipitation, widespread cold-season stratiform rain, and select tropical storm remnants, are observed and selected for analysis (Giangrande and Ryzhkov 2008). Concurrent gauge observations from the densely spaced Agricultural Research Service (ARS) and Oklahoma Mesonet (MES) network stations located 50–150 km (e.g., Fiebrich et al. 2006; McPherson et al. 2007; Shafer et al. 2000) from the KOUN radar are also included with this dataset.

JOBNAME: JTECH 00#0 2012 PAGE: 9 SESS: 8 OUTPUT: Wed Mar 14 20:58:42 2012 Total No. of Pages: 14
/ams/jtech/0/jtechD1100122

Journal of Atmospheric and Oceanic Technology (Proof Only)

Dual-polarized measurements (Z_H and Z_{DR}) from KOUN have been compared and calibrated using cross comparison with a disdrometer, the nearby KTLX radar (Oklahoma City, Oklahoma, WSR-88D), and polarimetric signatures of dry aggregated snow above the melting level. Attenuation correction in rain has been performed on Z_H and Z_{DR} using differential phase Φ_{dp}. Nonmeteorological echoes are filtered by a $\rho_{hv} > 0.85$ threshold. To mitigate hail contamination, the $Z_H < 53$ dBZ and $0 < Z_{DR} < 5$ dB thresholds were applied. Gauges further than 150 km from the radar have been removed to avoid–reduce partial beam filling and melting layer effects. Figures 7a,b show scatterplots of the Z_H and Z_{DR} measured at ARS and MES gauges. Compared to the scatterplots from the simulation, clear differences in distributions can be observed. There are extensive observations between 10 and 40 dBZ in the JPOLE dataset and the majority of the KOUN pairings have $Z_{DR} <$ 3 dB.

If hourly radar accumulations are defined as an hourly rainfall estimate centered on a gauge, then validation of GMRE can be performed by comparing hourly gauge and radar rainfall accumulations over gauge locations. Because usually only eight–nine radar scans are available over the same gauge location within 1 h, the nearest neighbor interpolation method is used to calculate hourly radar accumulations.

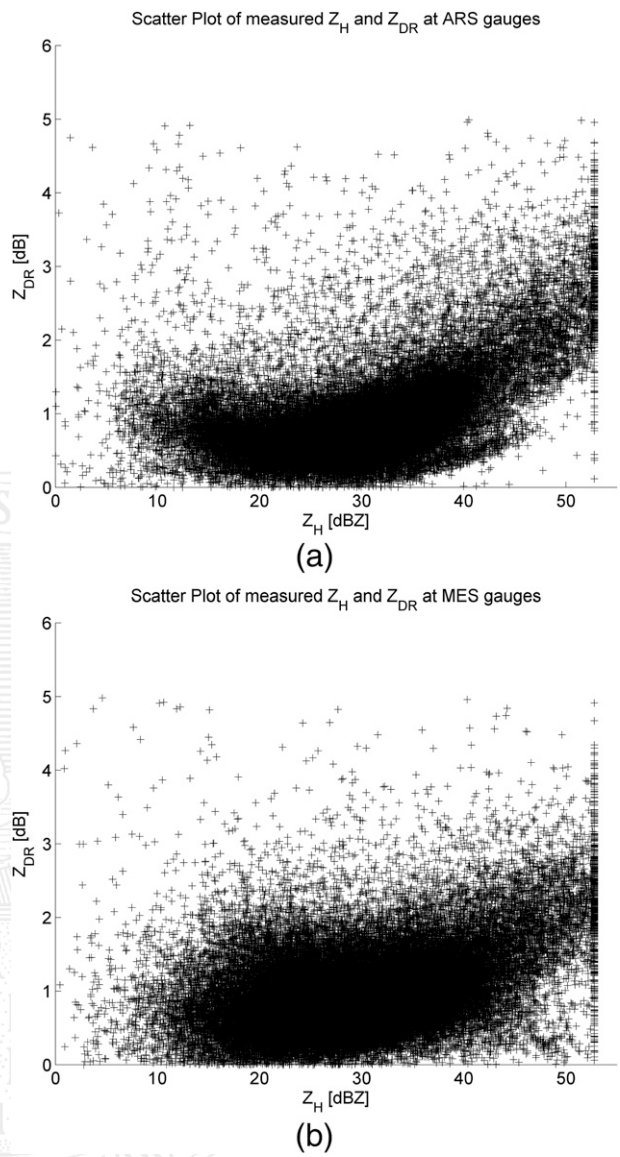

FIG. 7. Scatterplots of dataset-measured (a) Z_H and Z_{DR} at ARS gauges and (b) Z_H and Z_{DR} at MES gauges.

5. Results and comparisons

For a performance comparison of the GMRE approach, three rainfall-rate retrieval algorithms are selected. These relations are PLR form, based on an empirical regression of the measured gauge (or video disdrometer) and radar data. The first one, with Z_h as the only input, is the inversion of the standard NEXRAD rainfall formula for continental (nontropical) application (Fulton et al. 1998),

$$R(Z_h) = 1.7 \times 10^{-2} Z_h^{0.714}. \tag{18}$$

The second one, with Z_h and Z_{dr} as inputs, had optimized performance for rain in central Oklahoma during the JPOLE field campaign (Ryzhkov et al. 2005a),

$$R(Z_h, Z_{dr}) = 1.42 \times 10^{-2} Z_h^{0.77} Z_{dr}^{-1.67}. \tag{19}$$

The third, proposed in Ryzhkov et al. (2005b), combines the merits of different algorithms for various rain intensities and uses different combinations of radar variables Z_h, Z_{dr}, and K_{dp} based on rainfall rate estimated from Eq. (18),

$$R_{SYN} = \begin{cases} R(Z_h)/(0.4 + 5.0|Z_{dr} - 1|^{1.3}) & R(Z_h) < 6 \\ 44.0|K_{dp}|^{0.822}\mathrm{sign}(K_{dp})/(0.4 + 3.5|Z_{dr} - 1|^{1.7}) & 6 < R(Z_h) < 50 \\ 44.0|K_{dp}|^{0.822}\mathrm{sign}(K_{dp}) & R(Z_h) > 50 \end{cases}. \tag{20}$$

JOBNAME: JTECH 00#0 2012 PAGE: 10 SESS: 8 OUTPUT: Wed Mar 14 20:58:48 2012 Total No. of Pages: 14
/ams/jtech/0/jtechD1100122

TABLE 4. Performance comparison of rainfall retrieval algorithms for the ARS dataset. Unit: mm h^{-1}.

	$R(Z_h)$	$R(Z_h, Z_{dr})$	R_{SYN}	$R_{G5}(Z_H)$	$R_{G5}(Z_H, Z_{DR})$	R_{G5}	R_{G20}
Bias	1.68	−0.01	−0.24	−2.12	−0.80	−0.04	−0.49
STD	5.36	3.03	2.90	4.26	3.52	2.76	2.87
RMSE	5.62	3.03	2.91	4.76	3.61	2.76	2.91

TABLE 5. Performance comparison of rainfall retrieval algorithms for the MES dataset. Unit: mm h^{-1}.

	$R(Z_h)$	$R(Z_h, Z_{dr})$	R_{SYN}	$R_{G5}(Z_H)$	$R_{G5}(Z_H, Z_{DR})$	R_{G5}	R_{G20}
Bias	1.58	0.31	−0.3	−1.76	−0.57	0.19	−0.27
STD	5.24	4.44	3.12	4.56	4.05	2.98	3.30
RMSE	5.47	4.45	3.13	4.88	4.09	2.99	3.31

Note that the polarimetric algorithms in (19) and (20) have been optimized to perform well over the entire JPOLE dataset and were proven in later studies to be solid references for Oklahoma precipitation climatology. In comparison, because the GMRE is constructed from a simulation dataset generated from general microphysical parameterization, it should apply to and be capable in other precipitation regimes.

Two GMREs, one with 5 mixtures (G5) and the other with 20 mixtures (G20), are tested using this JPOLE dataset. Since noise properties of different dual-polarization variables in the JPOLE dataset are unknown, $\mathbf{R}$ is set to be zero in the current implementation. Because FSE statistics are heavily weighted toward small hourly precipitation accumulations, they are not examined during this test. Tables 4 and 5 summarize the results and comparisons of all retrieval algorithms over the ARS and MES gauges.

With reflectivity Z_H as input, $R_{G5}(Z_H)$ outperforms conventional NEXRAD $R(Z_h)$ in terms of RMSE for both datasets. With Z_H and Z_{DR} as inputs, $R_{G5}(Z_H, Z_{DR})$ performs slightly worse than the JPOLE $R(Z_h, Z_{dr})$ relation for the ARS dataset, but better for the MES dataset in terms of RMSE. With full polarimetric inputs Z_H, Z_{DR}, and K_{dp}, R_{G5} has the best performance in every category for both datasets; R_{G20} is comparable to R_{SYN} for the closer ARS dataset, but slightly worse than R_{SYN} for the MES dataset. All estimates but one from the GMREs show a negative bias, probably because of the fact that they are trained from a dataset that favors smaller rainfall. From the previous section, the GMRE with 20 mixtures converges closer to the distribution of the simulation dataset [denoted as $p_s(\mathbf{x})$], while the GMRE with 5 mixtures is only able to represent a general outline of $p_s(\mathbf{x})$ without many details. However, because $p_s(\mathbf{x})$ does not precisely match the distribution of the KOUN-based measurement dataset [denoted as $p_m(\mathbf{x})$], the GMRE with 20 mixtures is apparently overfitted to $p_s(\mathbf{x})$ and prohibits optimal performance in $p_m(\mathbf{x})$. However, the GMRE with five mixtures can outperform the JPOLE-tuned synthetic R_{SYN} relation in terms of bias and RMSE, even though it represents a less detailed $p_s(\mathbf{x})$, as highlighted in Figs. 8 and 9.

It is interesting to compare the performance of the simulation dataset–trained GMRE in this study with the neural network approach introduced in Vulpiani et al.

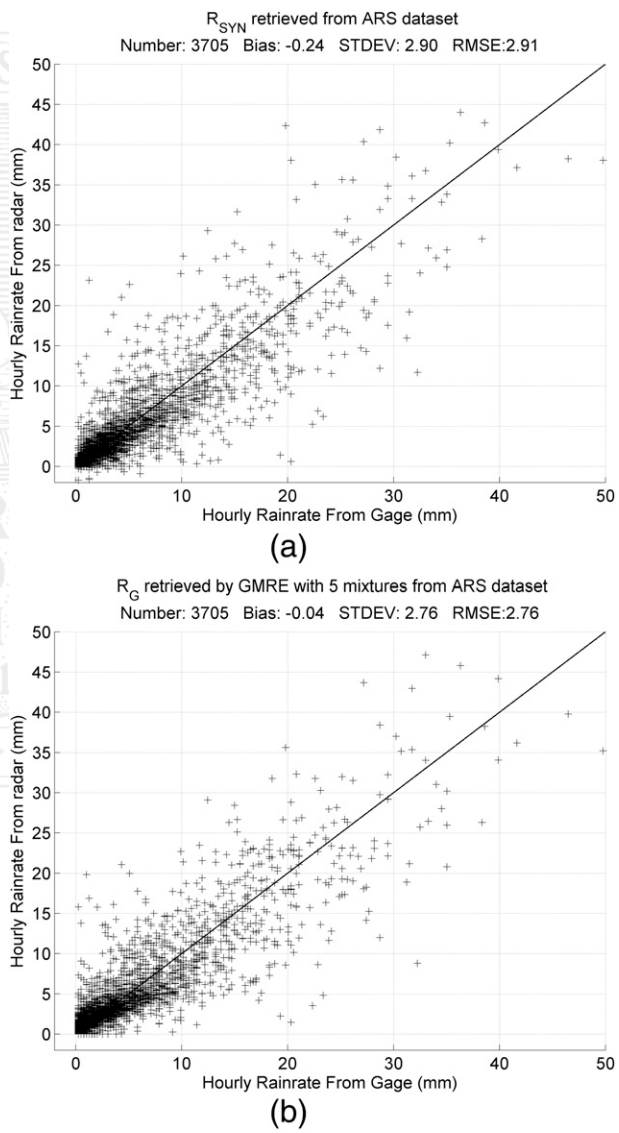

FIG. 8. Comparison plots of radar–gauge hourly accumulated rainfall rate for ARS dataset: (a) R_{SYN} retrieved from ARS dataset, and (b) R_G retrieved by GMRE with 5 mixtures from ARS dataset. Bias, STD, and RMSE are in millimeters for both plots.

Journal of Atmospheric and Oceanic Technology (Proof Only)

JOBNAME: JTECH 00#0 2012 PAGE: 11 SESS: 8 OUTPUT: Wed Mar 14 20:58:53 2012 Total No. of Pages: 14
/ams/jtech/0/jtechD1100122

Journal of Atmospheric and Oceanic Technology (Proof Only)

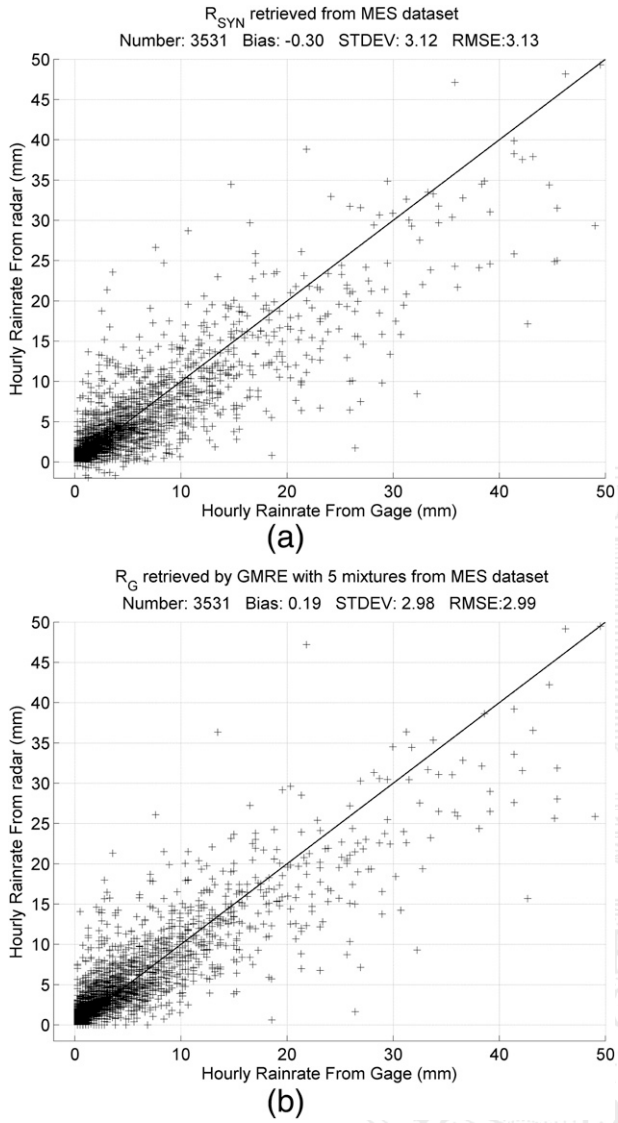

FIG. 9. Comparison plots of radar–gauge hourly accumulated rainfall rate for MES dataset: (a) R_{SYN} retrieved from MES dataset, and (b) R_G retrieved by GMRE with 5 mixtures from MES dataset. Bias, STD, and RMSE are in millimeters for both plots.

(2009) for the same ARS dataset. For these particular events [Table 2 in Vulpiani et al. (2009) and Table 4 herein], both the neural network and GMRE approach outperformed the synthetic relation, with the five-mixture GMRE showing a slightly better performance overall in terms of bias, STD, and RMSE.

These results confirmed that if GMRE is trained from a dataset whose distribution does not precisely match (but approximates) the distribution of the testing dataset, GMRE is still able to perform very well. When $p_s(\mathbf{x}) \neq p_m(\mathbf{x})$, the GMRE with less mixtures may perform even better than the GMRE with more mixtures, which could be overfitted to $p_s(\mathbf{x})$. However, depending on how $p_s(\mathbf{x})$ approximates $p_m(\mathbf{x})$ and how much $p_s(\mathbf{x})$ and $p_m(\mathbf{x})$ resemble one another, the optimal number of mixtures may vary. For example, comparing the dark blue area and light blue area where Z_H is between 15 and 50 dBZ in both Figs. 4a,b with the same areas in Figs. 7a,b, the distribution of G5 (Fig. 4a) at this area is clearly much closer to the same area of Figs. 7a,b than G20 (Fig. 4b). This explains why R_{G5} outperforms R_{G20} in both ARS and MES datasets. It also explains why extreme cases in the training dataset will not affect the performance of GMRE because only the relative probabilities of cases at areas of interest matter. As a consequence, GMRE should be trained from a large $p_s(\mathbf{x})$ that covers a broader range of occurrences than $p_m(\mathbf{x})$ (such as the extreme cases covered in the simulation of this study), to ensure that it is capable of handling not only a particular dataset, but also the radar observations from different seasons–regions. In addition to covering a broad range of scenarios, the training dataset also needs to have appropriate proportions of different kinds of precipitation. The more realistic the training dataset, the better the performance of GMRE can achieve in real environments. The performance of GMRE may be improved if it is trained from a dataset generated from measured DSDs over a wide range of time and area.

6. Discussion and conclusions

This study develops a Gaussian mixture rainfall-rate estimator for polarimetric radar–based rainfall-rate estimation. Theoretically, GMRE is the optimal estimator in terms of minimum variance and unbiased performance. It is also a general and flexible approach that can be adapted easily to different observation variables and rain types without compromising its performance. Training the GMRE is essential. During the training process, the prior distribution of microphysics parameters and observation variables is characterized and approximated by GMM, which converges to any specific distribution as the number of mixtures increases. In this study, the training dataset is constructed from a single-cell Monte Carlo simulation where the parameters of exponential DSD, N_0 and W, are randomly generated first from designed distributions that favor light and moderate rain. Then, uniformly distributed raindrops of different sizes are put into the single cell according to the DSD. Summing up waves scattered by each raindrop within the single cell, the Monte Carlo simulation produces realistic dual-polarization radar signatures.

GMREs with a different number of mixtures are trained and tested using the general simulation dataset. With more mixtures, the GMM converges more readily

JOBNAME: JTECH 00#0 2012 PAGE: 12 SESS: 8 OUTPUT: Wed Mar 14 20:58:58 2012 Total No. of Pages: 14
/ams/jtech/0/jtechD1100122

Journal of Atmospheric and Oceanic Technology (Proof Only)

to the simulation distribution, leading to better estimation. For the same radar observations, the rainfall rate directly estimated from the radar moments (R_G) is more precise than the rainfall rate retrieved from taking an indirect path through the estimated DSD parameters (R_{DSD}) wherein estimation error accumulates and magnifies.

Two GMREs, one with 5 mixtures and the other with 20 mixtures, in company with three PLR algorithms, are tested using the JPOLE dataset. As expected, better results are achieved when more radar observation variables are available for both the GMRE and PLR algorithms. While $R_{\mathrm{G5}}(Z_H, Z_{\mathrm{DR}})$ has a performance comparable to $R(Z_h, Z_{\mathrm{dr}})$, $R_{\mathrm{G5}}(Z_H)$ performs better than the single-parameter $R(Z_h)$ and R_{G5} outperforms the synthetic R_{SYN} JPOLE relation, which is the standard benchmark for this JPOLE dataset. However, R_{G20} does not perform as well as R_{G5}, which can be attributed to overfitting the GMRE to the specific simulation distribution that is dissimilar to the KOUN radar measurement distribution. Estimates from GMREs generally have a negative bias, which may reflect that these methods were trained from datasets that favor smaller rainfall over heavier rainfall, and also the fact that KOUN polarimetric radar inputs such as specific differential phase are smoothed somewhat in space–time.

In conclusion, GMRE shows great promise over conventional PLR techniques and provides a statistically optimized solution for rainfall-rate estimation. The convergence capability of GMM provides a general framework to accommodate extra information not only from dual-polarization diversities, but also from other diversities, such as multiple frequencies. A subject of ongoing research is to combine ground-based radar measurement with Ku–Ka band satellite radar measurements into the GMRE for better quantitative precipitation estimation (QPE). Because GMRE is a best estimator in terms of variance and bias performance, as long as the prior distribution is accurate, the focuses of rainfall-rate retrievals may be shifted from developing new algorithms–coefficients to constructing a better training dataset for GMRE. For example, better performance of GMRE may be achieved by tuning the distribution of N_0 and W in Monte Carlo simulations. If GMRE is trained from a dataset, either from simulation or measurement, without any climatologically driven optimization, then a global GMRE is possible for all rain types and regions. It is worth mentioning that applications of GMRE are not limited to the S band. Similarly, a GMRE can also be built for C- or X-band radars. Like other rainfall-rate estimation techniques, inputs to GMPE have to be corrected from attenuation before they could be used, especially in C and X band. As demonstrated in Li and Zhang (2011), attenuation correlations also can be incorporated into the GMRE framework, which will be studied in detail in future research.

Acknowledgments. The authors greatly appreciate the support from NOAA/NSSL to make this work possible, and the comments/suggestions of Dr. Alexander Ryzhkov from CIMMS/ARRC.

APPENDIX

Proof of Eq. (13)

Herein,

$$\begin{aligned}\mathcal{N}(\mathbf{z};\mathbf{Hx},\mathbf{R})\mathcal{N}(\mathbf{x};\boldsymbol{\mu}_i,\boldsymbol{\Sigma}_i) &= \frac{1}{\sqrt{(2\pi)}^{(m+n)}\sqrt{|\mathbf{R}|\cdot|\boldsymbol{\Sigma}_i|}}e^{-(1/2)||\mathbf{z}-\mathbf{Hx}||^2_{\mathbf{R}^{-1}}-(1/2)||\mathbf{x}-\boldsymbol{\mu}_i||^2_{\boldsymbol{\Sigma}_i^{-1}}}\\ &= \frac{1}{\sqrt{(2\pi)}^{(m+n)}\sqrt{C_2}}e^{-(1/2)C_1}\\ &= \mathcal{N}(\mathbf{z};\mathbf{H}\boldsymbol{\mu}_i,\mathbf{P}_i)\mathcal{N}(\mathbf{x};\hat{\boldsymbol{\mu}}_i,\hat{\mathbf{P}}_i)\end{aligned}$$

According to the Sherman–Morrison–Woodbury formula (Lewis et al. 2006),

$$\begin{aligned}\mathbf{P}_i^{-1} &= (\mathbf{H}\boldsymbol{\Sigma}_i\mathbf{H}^{\mathrm{T}}+\mathbf{R})^{-1}\\ &= \mathbf{R}^{-1}-\mathbf{R}^{-1}\mathbf{H}(\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}+\boldsymbol{\Sigma}_i^{-1})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\\ &= \mathbf{R}^{-1}-\mathbf{R}^{-1}\mathbf{H}\hat{\mathbf{P}}_i\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}.\end{aligned}$$

Then,

JOBNAME: JTECH 00#0 2012 PAGE: 13 SESS: 8 OUTPUT: Wed Mar 14 20:58:59 2012 Total No. of Pages: 14
/ams/jtech/0/jtechD1100122

$$
\begin{aligned}
C_1 &= \|\mathbf{x} - \hat{\boldsymbol{\mu}}_i\|^2_{\hat{\mathbf{P}}_i^{-1}} + \|\mathbf{z}\|^2_{\mathbf{P}_i^{-1}} - 2\boldsymbol{\mu}_i^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{P}_i^{-1}\mathbf{z} + \|\boldsymbol{\mu}_i\|^2_{\mathbf{H}^{\mathrm{T}}\mathbf{P}_i^{-1}\mathbf{H}} \\
&= \|\mathbf{x} - \hat{\boldsymbol{\mu}}_i\|^2_{\hat{\mathbf{P}}_i^{-1}} - 2\boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{\mathrm{T}}\hat{\mathbf{P}}_i\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{z} - \|\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{z}\|^2_{\hat{\mathbf{P}}_i} - \|\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\mu}_i\|^2_{\hat{\mathbf{P}}_i} + \|\mathbf{z}\|^2_{\mathbf{R}^{-1}} + \|\boldsymbol{\mu}_i\|^2_{\boldsymbol{\Sigma}_i^{-1}} \\
&= \|\mathbf{x} - \hat{\boldsymbol{\mu}}_i\|^2_{\hat{\mathbf{P}}_i^{-1}} - \|\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{z} + \boldsymbol{\Sigma}_i^{-1}\boldsymbol{\mu}_i\|^2_{\hat{\mathbf{P}}_i} + \|\mathbf{z}\|^2_{\mathbf{R}^{-1}} + \|\boldsymbol{\mu}_i\|^2_{\boldsymbol{\Sigma}_i^{-1}} \\
&= \|\mathbf{x} - \hat{\mathbf{P}}_i\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\mu}_i - \hat{\mathbf{P}}_i\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{z}\|^2_{\hat{\mathbf{P}}_i^{-1}} - \|\mathbf{H}^{\mathrm{T}}\mathrm{R}^{-1}\mathbf{z} + \boldsymbol{\Sigma}_i^{-1}\boldsymbol{\mu}_i\|^2_{\hat{\mathbf{P}}_i} + \|\mathbf{z}\|^2_{\mathbf{R}^{-1}} + \|\boldsymbol{\mu}_i\|^2_{\boldsymbol{\Sigma}_i^{-1}} \\
&= \|\mathbf{x}\|^2_{\boldsymbol{\Sigma}_i^{-1}+\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}} + \|\mathbf{z}\|^2_{\mathbf{R}^{-1}} + \|\boldsymbol{\mu}_i\|^2_{\boldsymbol{\Sigma}_i^{-1}} - 2\mathbf{z}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{x} - 2\boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\mathbf{x} \\
&= \|\mathbf{z} - \mathbf{H}\mathbf{x}\|^2_{\mathbf{R}-1} + \|\mathbf{x} - \boldsymbol{\mu}_i\|^2_{\boldsymbol{\Sigma}_i^{-1}}
\end{aligned}
$$

and

$$
\begin{aligned}
C_2 &= |\mathbf{R}| \cdot |\boldsymbol{\Sigma}_i| \\
&= \det\left(\begin{bmatrix} \mathbf{R} & 0 \\ 0 & \boldsymbol{\Sigma}_i \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} \mathbf{I} & 0 \\ \mathbf{H} & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{R} & 0 \\ 0 & \boldsymbol{\Sigma}_i \end{bmatrix}\begin{bmatrix} \mathbf{I} & \mathbf{H}^{\mathrm{T}} \\ 0 & \mathbf{I} \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} \boldsymbol{\Sigma}_i & \boldsymbol{\Sigma}_i\mathbf{H}^{\mathrm{T}} \\ \mathbf{H}\boldsymbol{\Sigma}_i & \mathbf{H}\boldsymbol{\Sigma}_i\mathbf{H}^{\mathrm{T}} + \mathbf{R} \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} \mathbf{I} & \mathbf{K}_i \\ 0 & \mathbf{I} \end{bmatrix}\begin{bmatrix} \hat{\mathbf{P}}_i & 0 \\ \mathbf{H}\boldsymbol{\Sigma}_i & \hat{\mathbf{P}}_i \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} \hat{\mathbf{P}}_i & 0 \\ 0 & \hat{\mathbf{P}}_i \end{bmatrix}\right) = |\hat{\mathbf{P}}_i| \cdot |\hat{\mathbf{P}}_i|.
\end{aligned}
$$

REFERENCES

Anagnostou, M. N., E. N. Anagnostou, and J. Vivekanandan, 2006: Correction for rain path specific and differential attenuation of X-band dual-polarization observations. *IEEE Trans. Geosci. Remote Sens.,* **44,** 2470–2480.

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.

Bringi, V. N., T. Tang, and V. Chandrasekar, 2004: Evaluation of a new polarimetrically based Z–R relation. *J. Atmos. Oceanic Technol.,* **21,** 612–623.

Cao, Q., G. Zhang, E. A. Brandes, and T. J. Schuur, 2010: Polarimetric radar rain estimation through retrieval of drop size distribution using a Bayesian approach. *J. Appl. Meteor. Climatol.,* **49,** 973–990.

Chiu, J. C., and G. W. Petty, 2006: Bayesian retrieval of complete posterior PDFs of oceanic rain rate from microwave observations. *J. Appl. Meteor. Climatol.,* **45,** 1073–1095.

Cifelli, R., V. Chandrasekar, S. Lim, P. C. Kennedy, Y. Wang, and S. A. Rutledge, 2011: A new dual-polarization radar rainfall algorithm: Application in Colorado precipitation events. *J. Atmos. Oceanic Technol.,* **28,** 352–364.

Di Michele, S., A. Tassa, A. Mugnai, F. S. Marzano, P. Bauer, and J. P. V. P. Baptista, 2005: Bayesian algorithm for microwave-based precipitation retrieval: Description and application to TMI measurements over ocean. *IEEE Trans. Geosci. Remote Sens.,* **43,** 778–791.

Doviak, R. J., J. K. Carter, V. M. Melnikov, and D. S. Zrnic, 2002: Modifications to the research WSR-88D to obtain polarimetric data. National Severe Storms Laboratory Rep., 49 pp. AU1

Evans, K. F., J. Turk, T. Wong, and G. L. Stephens, 1995: A Bayesian approach to microwave precipitation profile retrieval. *J. Appl. Meteor.,* **34,** 260–279.

Fiebrich, C. A., D. L. Grimsley, R. A. McPherson, K. A. Kesler, and G. R. Essenberg, 2006: The value of routine site visits in managing and maintaining quality data from the Oklahoma Mesonet. *J. Atmos. Oceanic Technol.,* **23,** 406–416.

Fulton, R. A., J. P. Breidenbach, D.-J. Seo, D. A. Miller, and T. O. Bannon, 1998: The WSR-88D rainfall algorithm. *Wea. Forecasting,* **13,** 377–395.

Giangrande, S. E., and A. V. Ryzhkov, 2008: Estimation of rainfall based on the results of polarimetric echo classification. *J. Appl. Meteor. Climatol.,* **47,** 2445–2462.

Gorgucci, E., V. Chandrasekar, V. N. Bringi, and G. Scarchilli, 2002: Estimation of raindrop size distribution parameters from polarimetric radar measurements. *J. Atmos. Sci.,* **59,** 2373–2384.

Hogan, R. J., 2007: A variational scheme for retrieving rainfall rate and hail reflectivity fraction from polarization radar. *J. Appl. Meteor. Climatol.,* **46,** 1544–1564.

Lee, G. W., 2006: Sources of errors in rainfall measurements by polarimetric radar: Variability of drop size distributions, observational noise, and variation of relationships between R and polarimetric parameters. *J. Atmos. Oceanic Technol.,* **23,** 1005–1028.

——, and I. Zawadzki, 2005: Variability of drop size distributions: Time-scale dependence of the variability and its effects on rain estimation. *J. Appl. Meteor.,* **44,** 241–255.

Lewis, J. M., S. Lakshmivarahan, and S. Dhall, 2006: *Dynamic Data Assimilation: A Least Squares Approach.* Vol. 104, *Encyclopedia of Mathematics and Its Applications,* Cambridge University Press, 680 pp.

Li, Z., and Y. Zhang, 2011: Application of Gaussian mixture model (GMM) and estimator to radar-based weather parameter estimations. *IEEE Geosci. Remote Sens. Lett.,* **8,** 1041–1045.

——, ——, G. Zhang, and K. A. Brewster, 2011: A microphysics-based simulator for advanced airborne weather radar development. *IEEE Trans. Geosci. Remote Sens.,* **49,** 1356–1373.

McPherson, R. A., and Coauthors, 2007: Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. *J. Atmos. Oceanic Technol.,* **24,** 301–321.

Journal of Atmospheric and Oceanic Technology (Proof Only)

JOBNAME: JTECH 00#0 2012 PAGE: 14 SESS: 8 OUTPUT: Wed Mar 14 20:58:59 2012 Total No. of Pages: 14
/ams/jtech/0/jtechD1100122

Mishchenko, M. I., 2000: Calculation of the amplitude matrix for a nonspherical particle in a fixed orientation. *Appl. Opt.,* **39,** 1026–1031.

Russell, S., and P. Norvig, 2009: *Artificial Intelligence: A Modern Approach.* 3rd ed. Prentice Hall, 1152 pp.

Ryzhkov, A. V., S. E. Giangrande, and T. J. Schuur, 2005a: Rainfall estimation with a polarimetric prototype of WSR-88D. *J. Appl. Meteor.,* **44,** 502–515.

——, T. J. Schuur, D. W. Burgess, P. L. Heinselman, S. E. Giangrande, and D. S. Zrnic, 2005b: The joint polarization experiment: Polarimetric rainfall measurements and hydrometeor classification. *Bull. Amer. Meteor. Soc.,* **86,** 809–824.

Seliga, T. A., and V. N. Bringi, 1976: Potential use of radar differential reflectivity measurements at orthogonal polarizations for measuring precipitation. *J. Appl. Meteor.,* **15,** 69–76.

Shafer, M. A., C. A. Fiebrich, D. S. Arndt, S. E. Fredrickson, and T. W. Hughes, 2000: Quality assurance procedures in the Oklahoma mesonetwork. *J. Atmos. Oceanic Technol.,* **17,** 474–494.

Vulpiani, G., F. S. Marzano, V. Chandrasekar, A. Berne, and R. Uijlenhoet, 2006: Polarimetric weather radar retrieval of raindrop size distribution by means of a regularized artificial neural network. *IEEE Trans. Geosci. Remote Sens.,* **44,** 3262–3275.

——, S. Giangrande, and F. S. Marzano, 2009: Rainfall estimation from polarimetric S-band radar measurements: Validation of a neural network approach. *J. Appl. Meteor. Climatol.,* **48,** 2022–2036.

Waldvogel, A., 1974: The N_0 jump of raindrop spectra. *J. Atmos. Sci.,* **31,** 1067–1078.

Wang, Y., and V. Chandrasekar, 2010: Quantitative precipitation estimation in the CASA X-band dual-polarization radar network. *J. Atmos. Oceanic Technol.,* **27,** 1665–1676.

Zhang, G., M. Xue, Q. Cao, and D. Dawson, 2008: Diagnosing the intercept parameter for exponential raindrop size distribution based on video disdrometer observations: Model development. *J. Appl. Meteor. Climatol.,* **47,** 2983–2992.

Journal of Atmospheric and Oceanic Technology (Proof Only)