**Ice particle type identification for shallow Arctic mixed-phase clouds using
X-band polarimetric radar**

Guang Wen¹

LACS, Institute of Atmospheric Physics, Chinese Academy of Sciences, China

Mariko Oue

The Pennsylvania State University, University Park, Pennsylvania, USA

Alain Protat

Centre for Australian Weather and Climate Research, Bureau of Meteorology,
Melbourne, Victoria, Australia

Johannes Verlinde

The Pennsylvania State University, University Park, Pennsylvania, USA

Hui Xiao

LACS, Institute of Atmospheric Physics, Chinese Academy of Sciences, China

¹ Corresponding author address: LACS, Institute of Atmospheric Physics, Chinese Academy of Sciences, Huayanli 40, Chaoyang District, Beijing 100029, China.
E-mail: wenguang1986@gmail.com

Abstract

Ice particle type identification for shallow Arctic mixed-phase clouds is studied using X-band polarimetric radar variables: horizontal reflectivity factor Z_h , differential reflectivity Z_{dr} , specific differential phase K_{dp} , and cross-correlation coefficient ρ_{hv} . The problem is formulated in a Bayesian classification framework, which consists of a probability density function (PDF) and a prior probability. The PDF is approximated using a Gaussian mixture model with parameters obtained by a clustering technique. The prior probability is constructed with the spatial contextual information based on a Markov random field. The PDF and prior probability are incorporated to produce the posterior probability, the maximum of which indicates the most likely particle type. The proposed algorithm is used to first derive the PDFs for the X-band polarimetric radar observations, and then identify the particle types within Arctic precipitating cloud cases sampled in Barrow, Alaska. The results are consistent with ground-based observations and the technique is capable of detecting and characterizing the variability of cloud microphysics in Arctic clouds.

Keywords:

Particle identification;
Polarimetric radar parameters
Microphysical characteristics
Arctic mixed-phase clouds

1. Introduction

The Arctic mixed-phase clouds, here defined as clouds composed of supercooled liquid drops and ice crystals within the same cloudy volume, often persist in the lower-troposphere for long periods, even up to a couple of weeks (Curry et al., 1996; Intrieri et al., 2002). The mixed-phase clouds are effective producers of various shapes of ice precipitating particles (i.e., sizes, aspect ratios, and densities), which affect radiative forcing from clouds and then contribute to the global and regional climate over the Arctic (Meehl and Washington, 1990; Houghton et al., 1991; Griggs and Noguer, 2002). The identification of the cloud particle types is important for understanding the microphysical and thermodynamic properties of the Arctic clouds.

State of the art polarimetric weather radars, transmitting and receiving horizontal and vertical linear polarizations separately, are sensitive to the microphysical properties of the cloud and precipitating particles, such as their compositions, effective sizes, and particle shapes (Bringi and Chandrasekar, 2001; Doviak and Znić, 2006). This allows for discrimination of a variety of particle types using the polarimetric radar measurements. Potential applications include improved quantitative precipitation estimation (Giangrande and Ryzhkov, 2008; Cifelli et al., 2011), cross validation between space-borne radars and ground-based polarimetric radars (Chandrasekar et al., 2008), and cloud microphysics research over the mid-latitudes (Vivekanandan et al., 1999). Recent case studies using X-band and Ka-band polarimetric radar measurements (Oue et al., 2015a; Oue et al., 2015b; Oue et al., 2015c) revealed ice precipitating particle growth in Arctic clouds (e.g., aggregation, depositional growth, riming, and secondary ice production). Oue et al. (2015a) showed larger differential reflectivity values suggesting light aggregation, light riming, and large pristine dendritic growth in Arctic shallow mixed-phase clouds than in mid-latitude snowstorms. In order to further analyze Arctic cloud microphysics and thermodynamics statistically, particle type identification (PID) algorithms are required.

Most of the PID algorithms adopted a fuzzy-logic based classification framework for mid-latitude summer clouds using S-band (e.g., Vivekanandan et al., 1999; Liu

and Chandrasekar, 2000; Straka et al., 2000; Zrnić et al., 2001; Schuur et al., 2003; Lim et al., 2005; Park et al., 2009; Chandrasekar et al., 2013), C-band (e.g., Keenan, 2003; May and Keenan, 2005; Marzano et al., 2007), and X-band (e.g., Dolan and Rutledge, 2009; He et al., 2010; Snyder et al., 2010; Liu et al., 2012) radar observations. In this approach, the allowed ranges of polarimetric variables are adjusted for each particle type by means of membership functions. The algorithms for winter storms were also investigated. Elmore (2011) evaluated the performance of the NSSL hydrometeor classification algorithm using S-band polarimetric data of cold seasons. Thompson et al. (2014) developed a winter storm classification algorithm, which has the capacity to discriminate dendrites, plates, dry aggregated snowflakes, rain, freezing rain, and sleet. In addition to the above fuzzy-logic methods, Marzano et al. (2008) and Li (2011) applied statistical decision theory to identification of particle types based on T-matrix backscattering simulations. For the specific case of Arctic clouds, which are characterized by complex and different ice growth processes relative to the mid-latitude storms, new scanning radar measurements have just been collected (Oue et al., 2015a; Oue et al., 2015b) requiring tailored PID techniques to be developed and evaluated.

Two difficulties hinder the application of the existing dual-polarization PID algorithms to the Arctic mixed-phase clouds. First, membership functions for the complex ice particle habits found in Arctic clouds are not available in the literature, therefore the existing fuzzy-logic classifiers cannot be readily updated with new classes in their identification schemes. Second, using radar backscattering simulations based on T-matrix or other techniques to derive those membership functions would require strong assumptions on particle shapes and particle size distributions for each specific ice particle habit, which may result in inaccurate identification.

In the present paper, our first objective is to characterize the probability density functions (PDFs) associated with the ice particle habits most commonly found in Arctic clouds using a clustering technique. The observational dataset was collected with a scanning dual-polarization X-band radar (X-SAPR) deployed by the US Department of Energy Atmospheric Radiation Measurement (ARM) program in

Barrow, Alaska. The clustering technique proposed by Wen et al. (2015) is employed to derive the PDFs for Arctic cloud particle types. The advantages of such a clustering approach are that: 1) the shape of the PDF is not pre-determined; 2) the identification of various particle types comes directly from the clustering technique rather than ad-hoc assumptions; and 3) it is an incremental method, in the sense that more prototypes can be added as more radar data become available.

The second objective of this paper is to present a PID scheme based on the work described in Wen et al. (2016) that makes use of these PDFs for various ice crystal types, but with additional constraints to ensure spatial consistency of the identification. Indeed, most PID techniques are pixel-based, which implies that the type at one range gate is independent from that of the surrounding gates. However, in natural clouds, the neighboring gates have a higher probability to contain the same particle type to form a homogeneous region where a given particle type is dominant. Therefore, the contextual information may be useful to improve the PID.

Section 2 introduces the Arctic precipitating radar measurements and in-situ ground observations. Section 3 describes the characterization of the PDFs of polarimetric variables associated with each ice particle habit. Section 4 presents a method for modeling the prior probability based on Markov random fields and implementation using iterative conditional modes. Section 5 illustrates experimental results for two contrasted Arctic shallow cloud cases: one primarily includes dendrites and aggregates of dendrites, and the other one mainly consists of irregular crystals and rimed dendrites. Finally, section 6 summarizes this paper.

2. Observational data

This study utilizes the radar data collected by the X-band scanning precipitation radar (X-SAPR) in Barrow, Alaska, in conjunction with precipitating particle photograph measurements during the North Slope of Alaska Radar Intensive Observing Period (Oue et al., 2015a, 1–7 May 2013, <http://www.arm.gov/campaigns/nsa2013nsasr>).

a. X-SAPR radar

The X-SAPR radar simultaneously transmits and receives horizontal and vertical polarizations and has the capacity of obtaining the 3-dimensional structures of Arctic clouds. This radar performed Hemispheric Range Height Indicator (HRHI) scans at azimuth angles of 7° , 52° , and 97° every 10 minutes and 17 Plan Position Indicator (PPI) scans between two adjacent HRHI scans during intensive observing periods between 1–7 May 2013. The HRHI scans sweep through elevation angles from 0° to 180° at each azimuth angle. Because polarimetric measurements have a strong dependence on the radar beam incident angle on the ice particles (e.g., Reinking et al., 1997; Wolde and Vali, 2001; Matrosov et al., 2012; Bechini et al., 2013), the analyses in this study used polarimetric measurements between elevation angles of $0-35^\circ$ and $145-180^\circ$. The PPI scans observed through 360° in azimuth at 16 elevation angles ($0-17.0^\circ$). Detailed case descriptions are presented in Oue et al. (2015a).

1) Polarimetric variables

The X-SAPR radar measures the polarimetric radar variables derived from the covariance matrix of the polarized radar returns. In this study, we use four polarimetric variables, including the horizontal reflectivity factor (Z_h), differential reflectivity (Z_{dr}), cross-correlation coefficient at zero lag (ρ_{hv}), and specific differential phase (K_{dp}).

The reflectivity factor Z_h (Z_v) at the horizontal (vertical) polarization is defined as (Bringi and Chandrasekar, 2001; Doviak and Zrnić, 2006)

$$Z_{h,v} = \frac{4\lambda^4}{\pi^4 |K|^2} \int_0^\infty |s_{hh,vv}(D)|^2 N(D) dD, \quad (1)$$

where D is the equivalent volume diameter, $N(D)$ is the drop size distribution in a resolution volume, K is the dielectric complex factor, λ is the radar wavelength, and $s_{hh,vv}$ is the amplitude of the horizontal (vertical) backscattering wave.

The differential reflectivity Z_{dr} is given by

$$Z_{dr} = 10 \log_{10} \left(\frac{Z_h}{Z_v} \right). \quad (2)$$

Z_{dr} is influenced by the axis ratio, density, size, and canting effects of particles (Seliga and Bringi, 1976).

The cross-correlation coefficient ρ_{hv} measures the degree of de-correlation between horizontally and vertically polarized signals. It is formulated as

$$\rho_{hv} = \frac{\int_0^{\infty} s_{hh}^*(D) s_{vv}(D) N(D) dD}{\left[\int_0^{\infty} |s_{hh}(D)|^2 N(D) dD \int_0^{\infty} |s_{vv}(D)|^2 N(D) dD \right]^{1/2}}. \quad (3)$$

The de-correlation results from the variability in the horizontal and vertical orientations of particles in a radar resolution volume.

The specific differential phase K_{dp} is a measure of the gradient of the local phase shifts due to anisotropic scatterers, because they produce different phase shifts for horizontally and vertically polarized waves. The definition can be given as

$$K_{dp} = \frac{180\lambda}{\pi} \left\{ \int_0^{\infty} \text{Re}[f_{hh}(D) - f_{vv}(D)] N(D) dD \right\}, \quad (4)$$

where $f_{hh}(D)$ and $f_{vv}(D)$ are the forward scattering amplitudes. K_{dp} is estimated from the range derivative of differential propagation phase (ϕ_{dp}) for horizontally and vertically polarized waves.

2) Quality control for X-SAPR data

The observed Z_{dr} values were corrected for a systematic bias of +0.94 dB by using zenith-pointing Z_{dr} calibration scans (Oue et al., 2015a; Oue et al., 2015b). Moreover, the Z_{dr} and ρ_{hv} values were corrected for noise biases, as described by Oue et al. (2015b). K_{dp} values were computed from ϕ_{dp} smoothed by the use of an iterative algorithm proposed by Hubbert and Bringi (1995). In the smoothing process, ϕ_{dp} values where $\rho_{hv} < 0.8$ are removed. If the difference between the filtered and raw

ϕ_{dp} values is greater than 2° , the filtered ϕ_{dp} value is subsequently used in the next iteration to obtain an updated ϕ_{dp} profile. K_{dp} is then computed as the slope of the smoothed ϕ_{dp} curve divided by 2; slopes are estimated using a least square method applied to 41 X-SAPR radar range gates over approximately 2 km of radial distance.

b. Photograph datasets

The particle characteristics at the ground are captured by photographs at a location approximately 0.3 km west of the X-SAPR radar. Table 1 shows the data and associated ice particle types at the ground, and Figure 1 gives examples of ice particle photographs observed on the ground. Five particle types can be identified from the X-SAPR and photograph measurements, including pristine dendrites (PR), aggregates of dendrites (AG), rimed dendrites (RD), spatial irregular crystals particles (IR), and plate-like crystals (PL).

3. Probability density functions

In this study, the probability density function (PDF) is modeled as a Gaussian mixture in a high-dimensional feature space composed by the polarimetric variables. The Gaussian mixture is commonly used to model the probability distributions for continuous variables because of its advantages over other mathematical models, those are: the ability to form a smooth approximation to any arbitrary shape of probability densities, the applicability to a complex multi-dimensional distribution characterized by a large number of Gaussian distributions, and the fast convergence to any specific distribution as the number of components grow. In this section, we apply a clustering technique to obtain the PDFs for the five particle types identified by the X-SAPR radar at the ground. These five of particle types are identified at the ground from photographs of surface precipitation particles. Note that the photographs are not used as inputs to the clustering technique.

a. Clustering method

In order to generate the PDFs of polarimetric variables in an objective way, we have used the similar clustering technique as that employed to classify precipitating hydrometeor types around Brisbane, Australia, with the CP-2 S-band radar (Wen et al., 2015). Figure 2 shows the flowchart of the clustering method used in this study. All details about the technique can be found in Wen et al. (2015), so in what follows we only briefly outline its main features. The technique uses a K -means clustering as a first step and an Expectation-Maximization (EM) clustering based on Gaussian Mixture Models (GMMs) as a second step. The EM clustering procedures are iteratively operated with initials for cluster centroids obtained by searching local minima and maxima in the data histograms. It generates a number of validated clusters, each of which associates with a particular particle type. The identification of the particle types of the clusters is achieved by comparing the regions of the clusters with the ground observations and previous analysis of the Arctic mixed-phase clouds in other studies (Oue et al., 2015a). The probability distributions of these clusters are promoted as prototypes, which are modeled as Gaussian distributions. The true PDF for each particle type is subsequently represented by as Gaussian mixture composed of a collection of these prototypes. In order to reduce the computation complexity caused by the increasing prototype number, a Kullback-Leibler divergence approach is implemented to merge the prototypes with similar means and covariances, but retain the original shape of the PDFs composed by all the prototypes. The prototypes with reduced number are used to classify the training data to check if they change the identification results. The PDFs obtained by the X-SAPR radar observations are studied and documented in order to study the polarimetric characteristics of the particles within the Arctic mixed-phase clouds.

b) Probability density functions for X-SAPR datasets

By applying the clustering technique to 117 RHI scans, we obtained 25 prototypes for pristine dendrites (PR), 15 for aggregates of dendrites (AG), 19 for rimed dendrites (RD), 24 for irregular crystals (IR), and 7 for plate-like crystals (PL).

Figure 3 shows the projections of the high-dimensional PR PDFs on selected

two-dimensional spaces (Z_h - Z_{dr} , Z_h - K_{dp} , and Z_h - ρ_{hv} , respectively). Z_h means are between -10.4 dBZ and 0.4 dBZ, while the Z_{dr} means are between 2.8 dB and 5.8 dB. The lower and upper values of K_{dp} means are -0.3 deg km⁻¹ and -0.0 deg km⁻¹. The means of ρ_{hv} are between 0.90 and 0.99. The projections of the PDFs for AG are illustrated in Figure 4. The means of Z_h are from -1.5 dBZ to 9.1 dBZ, Z_{dr} from 0.5 dB to 3.9 dB, K_{dp} from -0.5 deg km⁻¹ to 0.0 deg km⁻¹, and ρ_{hv} from 0.92 to 0.99. Figure 5 displays the projections of the RD PDFs composited from 19 prototypes. The means of Z_h is between -8.1 dBZ and 6.9 dBZ, while lower and upper values of Z_{dr} are 0.5 dB and 2.3 dB, respectively. The means of K_{dp} range from -0.2 deg km⁻¹ to 0.2 deg km⁻¹ and ρ_{hv} from 0.95 to 0.99. Figure 6 shows the projections of the IR PDFs derived from the 24 obtained prototypes. The lower and upper bounds of the Z_h means are -10.8 dBZ and 5.8 dBZ, respectively. The Z_{dr} ranges from -0.8 dB to 4.9 dB. Although the upper bound for Z_{dr} is 4.9 dB, most values are below 2 dB. The range of K_{dp} is between -0.6 deg km⁻¹ and 0.0 deg km⁻¹. The range of ρ_{hv} is between 0.96 and 0.99. Figure 7 gives the projections of the PDFs with 7 prototypes identified as plate-like crystals. The lower and upper limits of the Z_h means are -16.4 dBZ and -9.4 dBZ, respectively. The range of Z_{dr} is between 2.8 dB and 7.1 dB and K_{dp} between -0.2 deg km⁻¹ and -0.1 deg km⁻¹. The lower and upper bounds for the ρ_{hv} means are 0.95 and 1.08. A summary of means and standard deviations of the Gaussian mixture for each particle type is given in Table 2–Table 6. The proposed clustering technique and statistical quantities can also be used to construct the membership functions if a fuzzy logic method is applied. A list of thresholds derived from the clustering are shown in Table 7.

In order to highlight the differences between the sets of PDFs of polarimetric variables for each particle type and the discriminating power of the technique, Figure 8 shows PDFs for all particle types projected on the Z_h - Z_{dr} , Z_h - K_{dp} , and Z_h - ρ_{hv} two-dimensional spaces. The Z_h - Z_{dr} pair is the best discriminator for the five ice

particle types when compared with the Z_h - K_{dp} and Z_h - ρ_{hv} pairs. The most difficult particle types to discriminate are IR and RD, with IR particles only characterized by slightly larger Z_h and ρ_{hv} . K_{dp} does not appear to be useful to discriminate the five particle types, because there is not enough accumulation of differential phase shift with propagation to be detectable with reasonable range smoothing in these shallow Arctic clouds with often-low ice particle number concentrations and ice water content (Oue et al., 2015a). PR has a high Z_{dr} ($Z_{dr} > 2$ dB) with the maximal probability at around 5 dB, while AG has an intermediate Z_{dr} , between 1 dB and 3 dB. Such Z_{dr} values for aggregates in the present study are larger than those reported for aggregates in mid-latitude snowstorms, in which heavy aggregation occurred (e.g., Hall et al., 1984; Ryzhkov et al., 2005). The higher Z_{dr} values suggest the early stages of aggregation during which the aggregates retain more oblate shapes and higher density (Oue et al., 2015a). In contrast, RD and IR are characterized by low Z_{dr} with highest probability below 1 dB. PL often has a much lower Z_h than the other four types. The large range in Z_{dr} of PL reflects the natural variability in orientation of plate-like crystals. From the few prototypes we have, it seems that this particle type could be further split into horizontally-oriented and vertically oriented plates or to provide some indication of preferential orientation. However, we would need more prototypes to further refine the PDFs of polarimetric variables associated with these two orientations.

The PDFs are computed once and for all from the HRHI scans and stored in order to proceed with the classification. We can generally use the PDFs to conduct the particle identification or to construct the membership functions to form a fuzzy-logic classification (Wen et al., 2016). The classification results would be improved if a reasonable assumption can be made for modeling the prior probability. By combining the PDFs and prior probability, the problem is formulated in terms of Bayesian classification. This is described in the next section.

4. Bayesian method based on a Markov random field

This study proposes a particle identification (PID) algorithm based on the Bayesian classification method, which makes use of the PDFs of polarimetric variables for each particle type and the prior probability. In the previous study, we assume the prior probability is uniformly distributed (Wen et al., 2016). It is likely that the classification result would be improved if contextual information is used to model the prior probability. The use of context to help in the interpretation of weather radar imagery is a reasonable procedure, since a point classified as snow is likely to be surrounded by points of the same class and is unlikely to have neighbours from categories such as rain. The prior probability is constructed using the spatial inter-pixel dependency context in a local neighborhoods system. The Markov random field (MRF) approach is used to model the joint prior distribution of the range gate of interest and its surrounding neighbors. The PDF and prior probability are combined to yield the posterior probability. Finally, the particle type can be retrieved using a maximum a posterior probability inference.

a. Bayesian method

The Bayesian classification formula consists of two key components, namely the PDF and prior probability. The polarimetric radar data X is composed of n variables, i.e., $X_n(r)$, and $r = (i, j)$ denotes the index of a particular range gate, where i and j are the locations along the radial and elevation, respectively. Let C be a set of discrete values that represent the particle type of each gate in X . According to the Bayesian probability theorem, the law of the total probability states

$$\Pr(C | X) = \frac{\Pr(X | C) \cdot \Pr(C)}{\Pr(X)}, \quad (5)$$

where $\Pr(C | X)$ is defined as the posterior probability given the observed data vector X , $\Pr(X | C)$ is the class-conditional probability given that X is a member of type C , which is also referred to as the PDFs; $\Pr(C)$ is the prior probability, and $\Pr(X)$ is the probability of X . Since $\Pr(X)$ is independent of C , Equation (5) can be

simplified as

$$\Pr(C | X) \propto \Pr(C) \cdot \Pr(X / C). \quad (6)$$

Using these components in the Bayesian formula, the polarimetric radar data X is classified in terms of the Maximum-A-Posterior (MAP) probability inference to satisfy

$$\hat{C}_{\text{MAP}} = \arg \max_C \Pr(C | X) = \arg \max_C \{\Pr(X | C) \Pr(C)\}, \quad (7)$$

where $\hat{C}_{\text{MAP}}$ is an MAP estimate of class labels that maximizes the logarithm of the posterior probability. Note that the MAP estimate is pixel-based if the prior PDF have no contextual consequence in formulating Equation (7).

b. Prior probability

The prior probability can be modeled based on spatial contextual information within the data (Derin and Elliott, 1987; Dubes and Jain, 1989; Jhung and Swain, 1996; Solberg et al., 1996). In this approach the labels of the neighboring gates are treated as statistically dependent. In other word, a range gate is most likely to be the same type as the adjacent gates. The Markov random field approach (MRF, see appendix A for the definition) provides a convenient way to characterize the contextual information (Besag, 1974; Geman and Geman, 1984; Besag, 1986; Li, 1995).

According to the Hamersley-Clifford theorem, the prior probability $\Pr(C)$ can be expressed in terms of a Gibbs random field (GRF)

$$\Pr(C) = \frac{1}{Z} \exp \left[- \sum_q V_q(C) \right], \quad (8)$$

where Z is a normalizing factor, and $V_q(C)$ is an arbitrary function with regard to clique type q (see Appendix B for the definition of cliques), called the potential function. For simplicity of mathematics and computations, the potential functions are assumed to be restricted to the second-order neighborhoods system (see Appendix B for the definition of neighborhoods systems). The prior probability based on up to

pair-wise clique potential functions is expressed as

$$\Pr(C) = \frac{1}{Z} \exp[-V_{q_1}(C_r) - V_{q_2}(C_r, C_{N_r})], \quad (9)$$

where C_{N_r} is the labels of the neighbors of r in a clique q . The potential functions

$V_{q_1}(C_r)$ and $V_{q_2}(C_r, C_{N_r})$ are considered as

$$V_{q_1}(C_r) = \alpha \cdot C_r, \quad (10)$$

$$V_{q_2}(C_r, C_{N_r}) = \beta \cdot C_r \cdot C_{N_r}. \quad (11)$$

where α and β are constant coefficients.

For simplicity of the model, α is assumed to be zero for each gate r .

$V_{q_2}(C_{r_0}, C_{N_{r_0}})$, with regard to the second-order neighborhoods system centered at r_0 , is

defined as

$$V_{q_2}(C_{r_0}, C_{N_{r_0}}) = \beta \cdot B(C_{r_0}, C_{N_{r_0}}), \quad (12)$$

where

$$\beta = (\beta_1, \beta_2, \beta_3, \beta_4) \quad (13)$$

and

$$\begin{aligned} B(C_{r_0}, C_{N_{r_0}}) = & [B(C_{r_0}, C_{r_4}) + B(C_{r_0}, C_{r_8}), \\ & B(C_{r_0}, C_{r_2}) + B(C_{r_0}, C_{r_6}), \\ & B(C_{r_0}, C_{r_1}) + B(C_{r_0}, C_{r_5}), \\ & B(C_{r_0}, C_{r_3}) + B(C_{r_0}, C_{r_7})] \end{aligned} \quad (14)$$

$B(a, b)$ is a Boolean function:

$$B(a, b) = \begin{cases} 1 & \text{if } a = b \\ -1 & \text{otherwise} \end{cases}. \quad (15)$$

According to the equivalence of MRF and GRF (Besag, 1974; Kindermann and

Snell, 1980; see also in Appendix A), the conditional probability $\Pr(C_{r_0} | C_{N_{r_0}})$ can

be expressed as

$$\Pr(C_{r_0} | C_{N_{r_0}}) = \frac{\Pr(C_{r_0}, C_{N_{r_0}})}{\Pr(C_{N_{r_0}})} = \frac{\exp[-V_{q_2}(C_{r_0}, C_{N_{r_0}})]}{\sum_{C_{r_0}} \exp[-V_{q_2}(C_{r_0}, C_{N_{r_0}})]}. \quad (16)$$

Since the terms in the denominator in Equation (16) are independent of C_{r_0} , the numerator is also independent of C_{r_0} . If C_{r_0} is a different label for r_0 , then

$$\frac{\Pr(C_{r_0}, C_{N_{r_0}})}{\Pr(C'_{r_0}, C_{N_{r_0}})} = \frac{\exp\{-V_{q_2}(C_{r_0}, C_{N_{r_0}})\}}{\exp\{-V_{q_2}(C'_{r_0}, C_{N_{r_0}})\}}. \quad (17)$$

By substituting Equation (12) into Equation (17) and taking the logarithm of each side, β is given as

$$\beta = \frac{\ln \left\{ \frac{\Pr(C_{r_0}, C_{N_{r_0}})}{\Pr(C'_{r_0}, C_{N_{r_0}})} \right\}}{B(C'_{r_0}, C_{N_{r_0}}) - B(C_{r_0}, C_{N_{r_0}})}. \quad (18)$$

The joint probability $\Pr(C_{r_0}, C_{N_{r_0}})$ is estimated based on a histogram technique by counting the occurrence of a particular configuration of $(C_{r_0}, C_{N_{r_0}})$ (Dubes and Jain, 1989). Finally, the parameter vector β is solved by using a least-squares fit (Derin and Elliott, 1987).

c. Implementation

The Bayesian classification algorithm based on MRF is implemented using iterated conditional modes (ICM, Besag, 1986). This algorithm is feasible by updating the labels iteratively to avoid the computation of a large matrix. Two assumptions have been made to apply the ICM method. First, each gate is class-conditionally independent, so that

$$\Pr(X | C) = \prod_r \Pr(X_r | C_r). \quad (19)$$

Second, the labels of the dataset S are the realization of a MRF, and the probability mass functions are identical, i.e.,

$$\Pr(C_r | C'_r) = \Pr(C_r | C_{N_r}). \quad (20)$$

The basic procedures of the ICM method are as follows:

- 1) Initializing the configuration C_r by using the PDFs as derived in section 3;
- 2) Estimating the parameter vector β from the initial configuration C_r by using Equation (18);
- 3) Computing $\Pr(C_r | X_r) \propto \Pr(X_r | C_r) \times \Pr(C_r | C_{N_r})$ for every class C_r , $C_r \in L$, [Equation (6)], and update C_r by maximizing $\{\Pr(C_r | X_r), \forall C_r \in L\}$;
- 4) Repeating step 3) N times;
- 5) Select the label C_r with the maximal $\Pr(C_r | X_r)$ [Equation (7)].

The ICM finds the maximal posterior probability iteratively when the convergence criterion is satisfied. The experiments show that ten iterations are sufficient to ensure that the procedures converge. Due to the local property of the MRF, the algorithm can be implemented in a highly parallel manner.

5. Results and discussion

The X-SAPR observed five categories (pristine dendrites: PR, aggregates of dendrites: AG, rimed dendrites: RD, spatial irregular crystals particles: IR, and plate-like crystals: PL) of ice particle populations in two long-lasting precipitation events: a shallow boundary layer cloud system on 2 May 2013 and a frontal system on 6 May 2013.

a. 2 May 2013

The shallow boundary layer cloud system on 2 May 2013 was associated with a weak surface trough approaching Barrow, Alaska. It started with lightly precipitating, mixed-phase clouds with clouds top at 1.5 km. As the trough passed over Barrow multiple different cloud layers were identified and the precipitation intensified.

Photographs of particle types indicated that dendrites including small plate-like crystals dominated the surface precipitation from 1400 UTC to 1700 UTC, after which the surface precipitation type changed to aggregates of dendrites until 2050 UTC (Oue et al., 2015a).

Figure 9 shows the radar observations for an HRHI at 1611 UTC. The Z_{dr} reaches maximum values of up to 7 dB with low Z_h (< -5 dBZ) and high ρ_{hv} (> 0.97) values near the echo top, where PR is identified by the PID algorithm (Figure 10c). The Z_{dr} reduces with decreasing heights to about 2 – 3 dB near the surface (abscissa 5 – 22 km), with coincident increases in Z_h from -10 dBZ to 10 dBZ. This region is identified as AG by the PID algorithm (Figure 10c). These PID results are consistent with the detailed analysis reported in Oue et al. (2015a).

Figure 10 compares the new MRF-based Bayesian PID algorithm with the previous techniques developed for the subtropical summer storms, in which the PDFs of section 3 are introduced (Wen et al., 2016). Overall, the three techniques all provide very similar results, which is expected as they use the same prototypes or PDFs of polarimetric variables for each particle type. However, the Maximum Prototype Likelihood Classifier (MPLC, Figure 10a) and simple Bayesian classifier (Figure 10b) from Wen et al. (2015b) tend to classify small pockets of RD within the AG regions. These speckles are removed by using the prior probability based on Markov random field (Figure 10c), producing a smoother region of classification in comparison with the previous techniques.

b. 6 May 2015

On 6 May 2013 a deeper frontal system with multiple different liquid-cloud layers produced precipitation at Barrow with Z_h up to 12 dBZ. Photographs of particle types revealed that the surface precipitation was dominated by rimed dendrites between 0930 UTC and 1345 UTC, but switched to irregular crystals particles between 1600 UTC and 2040 UTC (Oue et al., 2015a).

Figure 11 shows an HRHI scan through this frontal system taken at 1836 UTC.

This scan exhibits signatures distinct from those of the shallow boundary layer cloud system on 2 May 2013. In contrast to the higher Z_{dr} measurements from the dendrites and aggregates of dendrites which dominated in 2 May case, the Z_{dr} on the 6 May case is generally below 2 dB. These low Z_{dr} values coexisted with relatively large Z_h (>2.5 dBZ), high ρ_{hv} (>0.97), and near-zero K_{dp} , where IR is identified based on the ground observations by photographs. Between the large Z_h regions (-15 to -10 dBZ), relatively higher Z_{dr} (2.5 dB to 3 dB) is associated with lower ρ_{hv} (<0.97), where RD is identified (Figure 12c). In the RD regions (Figure 12c), relatively large K_{dp} values (>0.2 ° km⁻¹) are found. These PID results are consistent (Oue et al., 2015a).

Figure 12 gives the comparison between the MPLC, simple Bayesian classifier, and the MRF-based Bayesian algorithms. Small speckles of aggregates of dendrites are detected in the left hemisphere in the MPLC (Figure 12a) and simple Bayesian classifier (Figure 12b), because the PDFs of aggregates of dendrites and rimed dendrites are very similar and difficult to delineate with the polarimetric variables, as discussed in section 3. Although the spatial contextual information allows the identified results to be corrected according to the types of their neighbours, dendrite patterns still exist in the results of the new PID algorithm. These may be produced by the scattered pixels of relatively high Z_{dr} ($Z_{dr} > 5$ dB).

As shown in Table 1, we used all HRHI scans as the training dataset to generate the PDFs, and PPI scans as the test dataset to evaluate the PID. The PID types in the PPI lowest elevation scans with no beam blockage are compared to the particle type dataset developed from photograph of the surface precipitation. The particle identification results for a PPI scan in the test dataset are shown in Figure 13. The result that irregular crystals dominates the particle types at the ground is consistent with the photographs. As shown in Figure 14, the comparison between the three methods confirms the new PID algorithm produces a more homogeneous region than the previous techniques, as expected.

c. Discussion

The notable feature of this study is the success of identifications for weak reflectivity regions (<10 dBZ), which generally dominate the Arctic precipitating clouds. The PDFs (Figure 8) show that the PDFs of Z_h versus Z_{dr} (Figure 8a) are effective discriminators for the ice particle types, while K_{dp} produced similar values around 0° km^{-1} for those types. In particular, RD is well separated from PR and PL by the Z_h versus Z_{dr} PDFs.

The HRHI scan at 1611 UTC 2 May 2013 (Figure 9 and Figure 10) documented an event dominated by the aggregation process of dendritic ice particles. (A comprehensive discussion this case can be found in Oue et al., 2015a Section 3a.) A thin liquid-cloud layer precipitated ice into an ice-supersaturated sub-cloud layer (Figures 4 and 5 in Oue et al., 2015a). Environmental temperatures between cloud top and surface varied from -13°C to -17°C , with two inversion layers present below the liquid-cloud layer. The top of the cloud (1.4 km) consists of a thin layer of PL. As these plate-like ice particles precipitate from the liquid-cloud layer they continue to grow by vapor deposition and transition to dendritic ice (PR). Aggregation starts at approximately 0.7 km above ground level, where the environmental temperature starts increasing from -17°C to -13°C at the surface. The onset of aggregation after below the 700 m deep layer of PR type ice crystals is consistent with the idea that the branches of dendritic crystals must be sufficiently developed before their edges will become entangled to form aggregates of dendrites (Houze, 1994). The three PID algorithms all indicate AG in this region, but the Bayesian method based on Markov random field modeling is most consistent with the conceptual models.

The HRHI scan at 1836 UTC 6 May 2013 (Figure 11 and Figure 12) shows a distinct microphysical process, which is primarily dominated by riming. On this day, the temperatures decreases from -18°C at cloud top to -5°C . The PID algorithms identify RD and IR, where RD corresponds to lower Z_h comparing to IR. However, negative Z_{dr} can be found at temperature of -12°C , -8°C , and -5°C . Relatively high Z_h (>5 dBZ) and near-zero Z_{dr} may be an indication of columnar or needle ice particles, but the PID algorithms classify them as IR. Further investigations need to be conducted for columns and needles. The PID results of the first elevation PPI scan

(Figure 13 and Figure 14) show a combination of RD and IR. This is consistent with the particle types documents with photographs of surface precipitation. Columnar or needle particles were not observed at the ground during this event.

In comparison with the existing PID methods, our categories are similar to those proposed in Matrosov et al. (2012) for observations of ice particles with a W-band radar, though they also investigated the columnar ice using slant linear depolarization ratio. Other PID methods often classify the three categories (PR, RD, and PL) into one category called “snow crystals” (e.g., Straka et al., 2000; Dolan and Rutledge, 2009). Crystals are often identified as snow, graupel, and sleet by the surface observations, which increase the error rate of the PID in the mid-latitude regions (Elmore, 2011). In contrast, our results are consistent with the particle types captured by the photographs at the ground. IR characterized as low Z_{dr} and high ρ_{hv} would correspond to the “ice crystals” category in Thompson et al. (2014). These finer categories can be helpful for quantitative estimations of ice water content and snowfall rate and validations of ice microphysics in numerical models capable of growing ice particles with diverse sizes, aspect ratios, and densities (e.g., Harrington et al., 2012; Sulia et al., 2013; Sulia et al., 2014).

The suitable application of context information based on Markov random fields allows the elimination of possible ambiguities, and the recovery of missing information. The result generated by the new PID algorithm show an improvement over the previous techniques in terms of smoothness of the classification region. The RD is effectively reduced within the AG zone (Figure 10), producing a homogeneous region of particles. However, the PR PDFs are very different from the PDFs for RD and AG. Therefore, the prior probability based on Markov random fields is most effective when the PDFs are close to each other and less effective when the likelihoods are well separated. The identification for an independent case and dataset has confirmed this conclusion (Figure 12 and Figure 14).

6. Conclusions

This paper proposed a particle identification (PID) algorithm to discriminate ice hydrometeors for Arctic clouds, which are commonly dominated by weak reflectivity (<10 dBZ), using a newly developed Bayesian method based on Markov random fields (MRF). The Bayesian method composes of the probability density function (PDF) and the prior probability. The PDFs of polarimetric radar variables for five different particle types (pristine dendrites, aggregates of dendrites, rimed dendrites, irregular ice and plate-like crystals) were constructed using a clustering method. The comparison between pairs of polarimetric variables shows that the PDFs of Z_h versus Z_{dr} are effective discriminators for the ice particle types, except for rimed dendrites and irregular ice crystals, which are characterized by similar PDFs of the polarimetric variables. The prior probability was modeled as a Markov random field (MRF) in order to consider the spatial contextual information within the data. The parameters in the MRF were estimated using a least square fit. The PDF and prior probability were combined to calculate the posterior probability. The Bayesian PID method based on MRF was finally approximated by iterated conditional modes.

The experimental results with Arctic precipitating clouds observed by X-band radar (X-SAPR) at Barrow, Alaska revealed the advantages of this PID algorithm. Two very different cases were analyzed, namely a shallow boundary layer cloud system on 2 May 2013 and a deeper frontal system on 6 May 2013. The clusters of polarimetric radar variables were classified into the five ice particles categories to generate prototypes, based on in-situ measurements at the ground. As with the method based on contextual information, it can improve the PID by reducing the speckle errors that are difficult to eliminate using the PDFs alone.

To identify other different hydrometeor species over the Arctic (e.g., raindrops, drizzle, melting particles, mixed-phase, and graupel), more case studies are needed, such as precipitation systems during summer including raindrops, drizzle, and melting particles; mixed-phase winter clouds; and deep winter cloud systems including graupel and heavy aggregation of particles. Recently, the Multi Angle Snowflake Camera (Garrett et al., 2012) has been installed at Oliktok, Alaska, collocated with scanning polarimetric radars. The combined in-situ and polarimetric radar datasets are

expected to reveal unique characteristics of polarimetric radar variables and their PDFs for other different hydrometeor species in the Arctic clouds.

Acknowledgement

This project is partially supported by the U.S. Department of Energy's Atmospheric Science Program Atmospheric System Research, an Office of Science, Office of Biological and Environmental Research program under Grant DE-FG02-05ER64058, the National Grand Fundamental Research 973 Programs of China (grant no. 2014CB441403, 2013CB430105), the National Natural Science Foundation of China (Grant Nos. 41575037, 41205099), Guizhou Province Scientific Research Joint Project (Grant No. G[2013]4001), and the Special Scientific Research Project of Meteorological Public Welfare Profession of China (Grant No. GYHY201006031), and China Postdoctoral Science Foundation (grant no. 2015M580123). We would like to express our sincere thanks to the anonymous reviewers for their valuable comments and suggestions.

Appendix A

Markov random field and equivalence of Markov random field and Gibbs random field

For a Markov random field $d = \{d_r\}_{r=1}^m$ defined on the set S , each variable d_r ($1 < r < m$) is assigned a label C from the label set L . The probability distribution of C satisfies

1) Positivity: $\Pr(C) > 0$ for all C ,

2) Markovianity: $\Pr(C_r | C'_r) = \Pr(C_r | C_{N_r})$, and

3) Homogeneity: $\Pr(C_r | C_{N_r})$ is the same for all r .

N_r is the neighborhoods system centered at gate r . C'_r denotes all the labels in S excluding at r .

According to the Hamersley-Clifford theorem, the prior probability $\Pr(C)$ can be expressed in terms of a Gibbs random field

$$\Pr(C) = \frac{1}{Z} \exp \left[- \sum_q V_q(C) \right]. \quad (21)$$

By the definition, the conditional probability $\Pr(C_r | C'_r)$ can be written as

$$\Pr(C_r | C'_r) = \frac{\Pr(C_r, C'_r)}{\Pr(C'_r)} = \frac{\Pr(C)}{\Pr(C'_r)}, \quad (22)$$

By substituting Equation (21), Equation (22) can be given as

$$\Pr(C_r | C'_r) = \frac{\exp \left[- \sum_q V_q(C) \right]}{\sum_{C'_r} \exp \left[- \sum_q V_q(C'_r) \right]}. \quad (23)$$

C consists of subset A and subset B , where A contains r and B excludes r . Equation (23) can be written as

$$\Pr(C_r | C'_r) = \frac{\exp \left[- \sum_{q \in A} V_q(C) \right] \times \exp \left[- \sum_{q \in B} V_q(C) \right]}{\sum_{C'_r} \exp \left[- \sum_{q \in A} V_q(C'_r) \right] \times \sum_{C'_r} \exp \left[- \sum_{q \in B} V_q(C'_r) \right]}. \quad (24)$$

Since $V_q(C) = V_q(C'_r)$ for subset B , Equation (24) can be simplified as

$$\Pr(C_r | C'_r) = \frac{\exp \left[- \sum_{q \in A} V_q(C) \right]}{\sum_{C'_r} \exp \left[- \sum_{q \in A} V_q(C'_r) \right]}. \quad (25)$$

If a second-order neighborhoods system is used, Equation (25) is reduced to

$$\Pr(C_r | C'_r) = \frac{\exp \left[- V_{q_2}(C_r, C_{N_r}) \right]}{\sum_{C'_r} \exp \left[- V_{q_2}(C'_r, C'_{N_r}) \right]}. \quad (26)$$

Equation (25) and Equation (26) indicate that the conditional probability

$\Pr(C_r | C'_r) = \Pr(C_r | C_{N_r})$ depends on the neighbors of r . Therefore, a Gibbs random

field is equivalent to a Markov random field as long as cliques are considered.

Appendix B

Neighborhoods systems and corresponding cliques

The first-order neighborhood system is defined as the four range gates sharing a side with a given gate, as shown in Figure_Apx 1a. The second-order neighborhood system is the eight gates having the same corner boundaries with the gate of interest, as illustrated in Figure_Apx 1b. Higher-order neighborhood systems can be defined by adding corners to a lower order neighborhood system.

A clique is a subset of the neighborhoods systems, so that all pairs of gates are mutual neighbors. The order of a clique is determined by the number of gates in the clique. Figure_Apx 2 illustrates different cliques of the second-order neighborhood system. The number of cliques grows as the order of neighborhoods system increases. For simplicity of mathematics and computations, we consider up to the pairwise clique in this study.

References

- Bechini, R., Baldini, L., Chandrasekar, V., 2013. Polarimetric Radar Observations in the Ice Region of Precipitating Clouds at C-Band and X-Band Radar Frequencies. *J. Appl. Meteor. Climatol.* 52, 1147-1169.
- Besag, J., 1974. Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society. Series B (Methodological)*. 36, 192-236.
- Besag, J., 1986. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society. B-48*, 259-302.
- Bringi, V.N., Chandrasekar, V., 2001. Polarimetric Doppler weather radar: Principles and applications. Cambridge Univ Pr.
- Chandrasekar, V., Bringi, V.N., Rutledge, S.A., Hou, A., Smith, E., Jackson, G.S., Gorgucci, E., Petersen, W.A., 2008. Potential Role Of Dual- Polarization Radar In The Validation Of Satellite Precipitation Measurements: Rationale and Opportunities. *Bull. Amer. Meteor. Soc.* 89, 1127-1145.
- Chandrasekar, V., Keranen, R., Lim, S., Moiseev, D., 2013. Recent advances in classification of observations from dual polarization weather radars. *Atmospheric Research*. 119, 97-111.
- Cifelli, R., Chandrasekar, V., Lim, S., Kennedy, P.C., Wang, Y., Rutledge, S.A., 2011. A new dual-polarization radar rainfall algorithm: Application in Colorado precipitation events. *J. Atmos. Oceanic Technol.* 28, 352-364.
- Curry, J.A., Schramm, J.L., Rossow, W.B., Randall, D., 1996. Overview of Arctic cloud and radiation characteristics. *J. Climate*. 9, 1731-1764.
- Derin, H., Elliott, H., 1987. Modeling and Segmentation of Noisy and Textured Images Using Gibbs Random Fields. *IEEE Trans. Pattern Anal. Mach. Intell. PAMI-9*, 39-55.
- Dolan, B., Rutledge, S.A., 2009. A theory-based hydrometeor identification algorithm for X-Band polarimetric radars. *J. Atmos. Oceanic Technol.* 26, 2071-2088.
- Doviak, R.J., Zrnić, D.S., 2006. Doppler radar and weather observations, 2nd ed. Dover Publications, Inc., Mineola, New York.

-
- 655 Dubes, R.C., Jain, A.K., 1989. Random field models in image analysis. *J Appl Stat.*
656 16, 131 - 164.
- 657 Elmore, K.L., 2011. The NSSL hydrometeor classification algorithm in winter surface
658 precipitation: Evaluation and future development. *Wea. Forecasting.* 26,
659 756-765.
- 660 Garrett, T.J., Fallgatter, C., Shkurko, K., Howlett, D., 2012. Fall speed measurement
661 and high-resolution multi-angle photography of hydrometeors in free fall.
662 *Atmos. Meas. Tech.* 5, 2625-2633.
- 663 Geman, S., Geman, D., 1984. Stochastic Relaxation, Gibbs Distributions, and the
664 Bayesian Restoration of Images. *IEEE Trans. Pattern Anal. Mach. Intell.*
665 *PAMI-6*, 721-741.
- 666 Giangrande, S.E., Ryzhkov, A.V., 2008. Estimation of rainfall based on the results of
667 polarimetric echo classification. *J. Appl. Meteor. Climatol.* 47, 2445-2462.
- 668 Griggs, D.J., Noguer, M., 2002. Climate change 2001: the scientific basis.
669 Contribution of working group I to the third assessment report of the
670 intergovernmental panel on climate change. *Weather.* 57, 267-269.
- 671 Hall, M.P.M., Goddard, J.W.F., Cherry, S.M., 1984. Identification of hydrometeors
672 and other targets by dual-polarization radar. *Radio Sci.* 19, 132-140.
- 673 Hallett, J., Mason, B., 1958. The influence of temperature and supersaturation on the
674 habit of ice crystals grown from the vapour. *Proceedings of the Royal Society*
675 *of London A: Mathematical, Physical and Engineering Sciences*, The Royal
676 Society, pp. 440-453.
- 677 Harrington, J.Y., Sulia, K., Morrison, H., 2012. A Method for Adaptive Habit
678 Prediction in Bulk Microphysical Models. Part I: Theoretical Development. *J.*
679 *Atmos. Sci.* 70, 349-364.
- 680 He, Y., Xiao, H., Lv, D., 2010. Analysis of Hydrometeor Distribution Characteristics
681 in Stratiform Clouds Using Polarization Radar. *Chinese Journal of*
682 *Atmospheric Sciences.* 34, 23-34. (In Chinese).
- 683 Houghton, D.D., Gallimore, R.G., Keller, L.M., 1991. Stability and variability in a
684 coupled ocean-atmosphere climate model: Results of 100-year simulations. *J.*

-
- 685 Climate. 4, 557-577.
- 686 Houze, R.A., 1994. Cloud dynamics. Academic Pr.
- 687 Hubbert, J., Bringi, V.N., 1995. An Iterative Filtering Technique for the Analysis of
- 688 Copolar Differential Phase and Dual-Frequency Radar Measurements. J.
- 689 Atmos. Oceanic Technol. 12, 643-648.
- 690 Intrieri, J.M., Fairall, C.W., Shupe, M.D., Persson, P.O.G., Andreas, E.L., Guest, P.S.,
- 691 Moritz, R.E., 2002. An annual cycle of Arctic surface cloud forcing at SHEBA.
- 692 Journal of Geophysical Research: Oceans. 107, SHE 13-11-SHE 13-14.
- 693 Jhung, Y., Swain, P.H., 1996. Bayesian contextual classification based on modified
- 694 M-estimates and Markov random fields. IEEE Trans. Geosci. Remote. 34,
- 695 67-75.
- 696 Keenan, T., 2003. Hydrometeor classification with a C-band polarimetric radar. Aust.
- 697 Meteor. Mag. 52, 23--31.
- 698 Kindermann, R., Snell, J.L., 1980. Markov Random Fields and Their Applications.
- 699 American Mathematical Society.
- 700 Li, S.Z., 1995. Markov random field modeling in computer vision. Springer-Verlag,
- 701 New York.
- 702 Li, Z., 2011. Applications of Gaussian mixture model to weather observations. PhD
- 703 Dissertation of The University of Oklahoma, United States, p. 207.
- 704 Lim, S., Chandrasekar, V., Bringi, V.N., 2005. Hydrometeor classification system
- 705 using dual-polarization radar measurements: Model improvements and in situ
- 706 verification. IEEE Trans. Geosci. Remote. 43, 792-801.
- 707 Liu, H., Chandrasekar, V., 2000. Classification of hydrometeors based on polarimetric
- 708 radar measurements: Development of fuzzy logic and neuro-fuzzy systems,
- 709 and in situ verification. J. Atmos. Oceanic Technol. 17, 140-164.
- 710 Liu, Y., Xiao, H., Yao, Z., Feng, L., 2012. Analyses of Hydrometeor Identification
- 711 Based on X-Band Polarimetric Radar. Climate and Environmental Research.
- 712 17, 925-936. (in Chinese)
- 713
- 714 Marzano, F.S., Scaranari, D., Montopoli, M., Vulpiani, G., 2008. Supervised

-
- 715 classification and estimation of hydrometeors from C-band dual-polarized
 716 radars: A Bayesian approach. *IEEE Trans. Geosci. Remote.* 46, 85-98.
- 717 Marzano, F.S., Scaranari, D., Vulpiani, G., 2007. Supervised fuzzy-logic
 718 classification of hydrometeors using C-band weather radars. *IEEE Trans.*
 719 *Geosci. Remote.* 45, 3784-3799.
- 720 Matrosov, S.Y., Mace, G.G., Marchand, R., Shupe, M.D., Hallar, A.G., McCubbin,
 721 I.B., 2012. Observations of Ice Crystal Habits with a Scanning Polarimetric
 722 W-Band Radar at Slant Linear Depolarization Ratio Mode. *J. Atmos. Oceanic*
 723 *Technol.* 29, 989-1008.
- 724 May, P.T., Keenan, T.D., 2005. Evaluation of microphysical retrievals from
 725 polarimetric radar with wind profiler data. *J. Appl. Meteor.* 44, 827-838.
- 726 Meehl, G.A., Washington, W.M., 1990. CO2 climate sensitivity and snow-sea-ice
 727 albedo parameterization in an atmospheric GCM coupled to a mixed-layer
 728 ocean model. *Clim Change.* 16, 283-306.
- 729 Oue, M., Galletti, M., Verlinde, J., Ryzhkov, A., Lu, Y., 2015a. Use of X-band
 730 differential reflectivity measurements to study shallow Arctic mixed-phase
 731 clouds. in press.
- 732 Oue, M., Kumjian, M.R., Lu, Y., Jiang, Z., Clothiaux, E.E., Verlinde, J., Aydin, K.,
 733 2015b. X-Band Polarimetric and Ka-Band Doppler Spectral Radar
 734 Observations of a Graupel-Producing Arctic Mixed-Phase Cloud. *J. Appl.*
 735 *Meteor. Climatol.* 54, 1335-1351.
- 736 Oue, M., Kumjian, M.R., Lu, Y., Verlinde, J., Aydin, K., Clothiaux, E.E., 2015c.
 737 Linear Depolarization Ratios of Columnar Ice Crystals in a Deep Precipitating
 738 System over the Arctic Observed by Zenith-Pointing Ka-Band Doppler Radar.
 739 *J. Appl. Meteor. Climatol.* 54, 1060-1068.
- 740 Park, H.S., Ryzhkov, A.V., Zrnić, D.S., Kim, K.-E., 2009. The hydrometeor
 741 classification algorithm for the polarimetric WSR-88D: Description and
 742 application to an MCS. *Wea. Forecasting.* 24, 730-748.
- 743 Reinking, R.F., Matrosov, S.Y., Brientjes, R.T., Martner, B.E., 1997. Identification of
 744 Hydrometeors with Elliptical and Linear Polarization Ka-Band Radar. *J. Appl.*

-
- 745 Meteor. 36, 322-339.
- 746 Ryzhkov, A.V., Giangrande, S.E., Melnikov, V.M., Schuur, T.J., 2005. Calibration
 747 Issues of Dual-Polarization Radar Measurements. J. Atmos. Oceanic Technol.
 748 22, 1138-1155.
- 749 Schuur, T., Ryzhkov, A., Heinselman, P., 2003. Observation and classification of
 750 echoes with the polarimetric WSR-88D radar. NOAA/NSSL Rep., p. 46.
- 751 Seliga, T.A., Bringi, V.N., 1976. Potential use of radar differential reflectivity
 752 measurements at orthogonal polarizations for measuring precipitation. J. Appl.
 753 Meteor. 15, 69-76.
- 754 Snyder, J.C., Bluestein, H.B., Zhang, G., Frasier, S.J., 2010. Attenuation correction
 755 and hydrometeor classification of high-resolution, X-band, dual-polarized
 756 mobile radar measurements in severe convective storms. J. Atmos. Oceanic
 757 Technol. 27, 1979-2001.
- 758 Solberg, A.H.S., Taxt, T., Jain, A.K., 1996. A Markov random field model for
 759 classification of multisource satellite imagery. IEEE Trans. Geosci. Remote.
 760 34, 100-113.
- 761 Straka, J.M., Zrnić, D.S., Ryzhkov, A.V., 2000. Bulk hydrometeor classification and
 762 quantification using polarimetric radar data: Synthesis of relations. J. Appl.
 763 Meteor. 39, 1341-1372.
- 764 Sulia, K.J., Harrington, J.Y., Morrison, H., 2013. A Method for Adaptive Habit
 765 Prediction in Bulk Microphysical Models. Part III: Applications and Studies
 766 within a Two-Dimensional Kinematic Model. J. Atmos. Sci. 70, 3302-3320.
- 767 Sulia, K.J., Morrison, H., Harrington, J.Y., 2014. Dynamical and Microphysical
 768 Evolution during Mixed-Phase Cloud Glaciation Simulated Using the Bulk
 769 Adaptive Habit Prediction Model. J. Atmos. Sci. 71, 4158-4180.
- 770 Thompson, E.J., Rutledge, S.A., Dolan, B., Chandrasekar, V., Cheong, B.L., 2014. A
 771 Dual-Polarization Radar Hydrometeor Classification Algorithm for Winter
 772 Precipitation. J. Atmos. Oceanic Technol. 31, 1457-1481.
- 773 Vivekanandan, J., Ellis, S.M., Oye, R., Zrnić, D.S., Ryzhkov, A.V., Straka, J., 1999.
 774 Cloud microphysics retrieval using S-band dual-polarization radar

-
- 775 measurements. Bull. Amer. Meteor. Soc. 80, 381-388.
- 776 Wen, G., Protat, A., May, P.T., Moran, W., Dixon, M., 2016. A Cluster-Based
- 777 Method for Hydrometeor Classification Using Polarimetric Variables. Part II:
- 778 Classification. J. Atmos. Oceanic Technol. In press.
- 779 Wen, G., Protat, A., May, P.T., Wang, X., Moran, W., 2015. A Cluster-Based Method
- 780 for Hydrometeor Classification Using Polarimetric Variables. Part I:
- 781 Interpretation and Analysis. J. Atmos. Oceanic Technol. 32, 1320-1340.
- 782 Wolde, M., Vali, G., 2001. Polarimetric Signatures from Ice Crystals Observed at 95
- 783 GHz in Winter Clouds. Part I: Dependence on Crystal Form. J. Atmos. Sci. 58,
- 784 828-841.
- 785 Zrnić, D.S., Ryzhkov, A., Straka, J., Liu, Y., Vivekanandan, J., 2001. Testing a
- 786 procedure for automatic classification of hydrometeor types. J. Atmos.
- 787 Oceanic Technol. 18, 892-913.
- 788

789

		Training dataset	Test dataset
Ground Type	Date (dd/mm/year)	HRHI (Time UTC)	PPI (Time UTC)
Pristine dendrites (PR)	02/05/2013	1415, 1420, 1425, 1430, 1431, 1436, 1441, 1512, 1609, 1611, 1616, 1644	
Aggregates of dendrites (AG)	02/05/2013	1425, 1430, 1609, 1611, 1616, 1724, 1727, 1838	1927, 1939
Rimed dendrites (RD)	06/05/2013	0901, 0906, 1057, 1102, 1107, 1114	
Irregular crystals (IR)	06/05/2013	1822, 1827, 1836, 1841, 1908, 1913, 1922, 1927, 1936, 1941, 1950, 1955	1816, 1830, 1901, 1915, 1930

790 Table 1 Training and test Datasets for particle identification. 117 HRHI scans are
791 applied for analysis of PDFs, and 111 PPI scans are used to test the particle
792 identification algorithm.

793

i	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$	$i=9$	$i=10$
w	0.007	0.001	0.006	0.001	0.013	0.012	0.014	0.003	0.004	0.020
μ_1	-4.8	-1.8	-6.0	-8.7	-2.8	-4.2	-3.1	-10.4	-1.7	-3.1
μ_2	5.5	4.6	5.4	5.0	5.0	4.9	5.2	4.3	4.9	5.0
μ_3	-0.2	-0.1	-0.1	-0.1	-0.0	-0.2	-0.1	-0.1	-0.1	-0.1
μ_4	0.97	0.99	0.98	0.94	0.96	0.99	0.98	0.90	0.94	0.98
σ_1	3.29	2.83	5.24	5.43	4.92	6.22	5.46	6.86	3.80	5.99
σ_2	0.95	0.85	1.29	1.22	1.35	1.30	1.37	3.13	1.25	1.42
σ_3	0.54	0.42	0.43	0.40	0.50	0.50	0.40	0.54	0.41	0.44
σ_4	0.022	0.004	0.017	0.017	0.030	0.008	0.013	0.139	0.017	0.013
i	$i=11$	$i=12$	$i=13$	$i=14$	$i=15$	$i=16$	$i=17$	$i=18$	$i=19$	$i=20$
w	0.015	0.009	0.019	0.021	0.011	0.008	0.017	0.011	0.012	0.011
μ_1	-7.2	-6.2	-3.9	-2.0	-0.8	0.4	-8.1	-1.4	-1.0	-7.3
μ_2	3.8	5.4	4.6	5.2	5.2	4.8	5.6	2.8	4.5	5.8
μ_3	-0.1	-0.1	-0.2	-0.3	-0.0	-0.2	-0.1	-0.1	-0.1	-0.1
μ_4	0.94	0.98	0.94	0.98	0.95	0.98	0.97	0.93	0.98	0.98
σ_1	5.69	4.70	5.39	4.34	4.85	5.67	6.14	4.14	3.37	5.49
σ_2	2.44	1.32	2.04	1.24	1.48	1.53	1.25	2.01	1.34	1.20
σ_3	0.39	0.38	0.66	0.63	0.57	0.71	0.53	0.41	0.45	0.43
σ_4	0.051	0.020	0.050	0.013	0.037	0.011	0.028	0.039	0.008	0.024
i	$i=21$	$i=22$	$i=23$	$i=24$	$i=25$					
w	0.021	0.014	0.015	0.015	0.010					
μ_1	-5.3	-3.8	-4.3	-0.6	-1.7					

μ_2	5.4	5.0	5.7	4.7	5.1
μ_3	-0.1	-0.1	-0.1	-0.0	-0.0
μ_4	0.97	0.98	0.98	0.95	0.94
σ_1	5.69	5.79	4.27	4.31	4.59
σ_2	1.26	1.39	1.14	1.44	1.48
σ_3	0.38	0.36	0.39	0.38	0.39
σ_4	0.024	0.011	0.014	0.022	0.031

794 Table 2 Components of a Gaussian mixture model (GMM) for PR. i is the component
795 index of GMM. w is the corresponding weight for the component. μ_1, μ_2, μ_3 , and
796 μ_4 are the means of Z_h (dBZ), Z_{dr} (dB), K_{dp} (deg km⁻¹), and ρ_{hw} , respectively. σ_1, σ_2 ,
797 σ_3 and σ_4 are the standard deviations.

798

i	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$	$i=9$	$i=10$
w	0.018	0.006	0.011	0.012	0.010	0.023	0.017	0.004	0.007	0.007
μ_1	1.9	9.1	7.2	4.7	5.4	4.6	6.1	0.2	7.4	6.3
μ_2	3.3	2.6	2.5	3.5	3.0	3.6	2.6	0.5	3.3	2.7
μ_3	-0.2	-0.5	-0.2	-0.2	-0.0	-0.0	-0.3	-0.0	0.0	-0.0
μ_4	0.95	0.99	0.97	0.98	0.96	0.96	0.98	0.96	0.97	0.97
σ_1	3.71	2.31	1.88	3.66	2.49	2.91	3.21	3.84	3.32	2.51
σ_2	1.57	0.79	0.73	1.11	0.94	1.15	1.24	0.70	1.03	0.56
σ_3	0.58	0.55	0.49	0.68	0.30	0.55	0.65	0.26	0.38	0.28
σ_4	0.026	0.007	0.010	0.011	0.014	0.021	0.011	0.006	0.019	0.014
i	$i=11$	$i=12$	$i=13$	$i=14$	$i=15$					
w	0.006	0.008	0.015	0.017	0.008					
μ_1	4.7	0.97	-1.5	6.8	1.7					
μ_2	3.9	3.2	3.4	2.8	3.9					
μ_3	-0.1	-0.1	-0.1	-0.3	-0.1					
μ_4	0.97	0.98	0.92	0.97	0.99					
σ_1	2.96	2.73	4.75	2.87	3.24					
σ_2	0.74	1.26	2.02	0.86	1.18					
σ_3	0.27	0.40	0.57	0.53	0.57					
σ_4	0.013	0.009	0.039	0.017	0.007					

Table 3 Same as Table 2, but for AG.

i	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$	$i=9$	$i=10$
w	0.023	0.018	0.015	0.010	0.012	0.020	0.003	0.003	0.008	0.012
μ_1	-8.1	-4.2	-4.3	-7.9	-6.3	-2.2	6.9	4.2	-2.2	-5.2
μ_2	1.0	0.5	0.8	0.9	0.7	2.3	1.6	2.0	1.6	1.3
μ_3	-0.1	-0.1	-0.1	-0.1	-0.0	-0.1	-0.1	-0.1	-0.0	-0.1
μ_4	0.97	0.98	0.99	0.98	0.97	0.95	0.98	0.96	0.96	0.96
σ_1	5.07	4.74	4.66	3.67	4.72	5.04	1.44	1.77	3.64	6.66
σ_2	1.06	0.63	0.72	1.07	0.99	1.74	0.80	0.61	0.90	1.24
σ_3	0.32	0.25	0.27	0.25	0.48	0.54	0.22	0.23	0.31	0.36
σ_4	0.025	0.010	0.005	0.013	0.034	0.037	0.005	0.012	0.022	0.032
i	$i=11$	$i=12$	$i=13$	$i=14$	$i=15$	$i=16$	$i=17$	$i=18$	$i=19$	
w	0.004	0.009	0.013	0.022	0.010	0.019	0.012	0.004	0.018	
μ_1	-0.5	-0.2	-2.2	3.5	-0.8	-1.7	5.9	-6.4	-7.3	
μ_2	1.4	1.6	1.9	0.8	1.7	1.3	1.4	0.8	1.3	
μ_3	-0.1	-0.0	-0.0	-0.1	-0.1	-0.1	-0.1	0.2	-0.2	
μ_4	0.99	0.96	0.95	0.98	0.96	0.97	0.97	0.95	0.98	
σ_1	1.87	3.25	4.50	3.57	3.80	4.17	2.60	4.77	4.97	
σ_2	0.79	0.84	1.56	0.69	0.88	1.03	0.92	0.80	0.82	
σ_3	0.36	0.26	0.44	0.25	0.26	0.43	0.24	0.31	0.23	
σ_4	0.009	0.014	0.036	0.010	0.017	0.021	0.012	0.022	0.011	

803 Table 4 Same as Table 2, but for RD.

804

i	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$	$i=9$	$i=10$
w	0.011	0.009	0.023	0.002	0.011	0.016	0.022	0.014	0.019	0.028
μ_1	-2.3	-0.5	-1.3	1.4	0.6	-2.7	1.1	4.3	-0.6	0.7
μ_2	0.6	0.5	0.5	2.4	0.5	0.7	0.7	0.9	1.1	0.7
μ_3	-0.2	-0.2	-0.1	-0.0	-0.3	-0.1	-0.1	-0.2	-0.1	-0.1
μ_4	0.99	0.98	0.99	0.98	0.99	0.98	0.99	0.99	0.98	0.99
σ_1	4.26	3.57	4.63	5.52	3.91	4.30	4.73	3.60	4.12	6.02
σ_2	0.66	0.65	0.62	1.43	0.60	0.89	0.67	0.80	0.90	0.66
σ_3	0.45	0.46	0.27	0.86	0.48	0.42	0.29	0.46	0.38	0.26
σ_4	0.005	0.011	0.009	0.010	0.007	0.015	0.008	0.006	0.015	0.008
i	$i=11$	$i=12$	$i=13$	$i=14$	$i=15$	$i=16$	$i=17$	$i=18$	$i=19$	$i=20$
w	0.013	0.006	0.004	0.023	0.009	0.011	0.010	0.000	0.012	0.003
μ_1	5.8	-7.4	-1.9	4.2	-4.4	4.0	4.5	0.8	-5.3	-1.5
μ_2	0.7	0.9	0.4	0.8	0.9	0.9	0.7	-0.8	1.2	4.9
μ_3	-0.2	-0.1	0.0	-0.1	0.0	-0.2	-0.1	-0.1	0.0	-0.6
μ_4	0.99	0.99	0.99	0.98	0.97	0.99	0.98	0.97	0.97	0.99
σ_1	3.14	3.18	3.24	3.14	4.38	3.98	3.17	2.19	4.45	3.57
σ_2	0.69	0.80	0.94	0.67	0.96	0.76	0.60	0.66	1.00	1.62
σ_3	0.47	0.29	0.43	0.26	0.31	0.48	0.39	0.44	0.28	0.37
σ_4	0.007	0.009	0.011	0.009	0.020	0.006	0.012	0.007	0.023	0.009
i	$i=21$	$i=22$	$i=23$	$i=24$						
w	0.016	0.017	0.003	0.004						
μ_1	3.3	-3.5	-10.8	-3.6						

μ_2	1.1	0.5	-0.2	0.3
μ_3	-0.2	-0.1	0.0	-0.1
μ_4	0.99	0.99	0.96	0.98
σ_1	3.34	4.51	3.14	4.53
σ_2	0.85	0.61	0.81	0.58
σ_3	0.48	0.26	0.32	0.23
σ_4	0.007	0.006	0.020	0.003

805 Table 5 Same as Table 2, but for IR.

806

<i>i</i>	<i>i</i> =1	<i>i</i> =2	<i>i</i> =3	<i>i</i> =4	<i>i</i> =5	<i>i</i> =6	<i>i</i> =7
<i>w</i>	0.002	0.002	0.002	0.010	0.000	0.001	0.009
μ_1	-12.0	-14.6	-16.4	-9.4	-15.3	-12.9	-12.6
μ_2	5.1	2.9	5.7	5.4	7.1	5.1	2.8
μ_3	-0.2	-0.1	-0.2	-0.1	-0.1	-0.1	-0.1
μ_4	1.00	0.96	1.00	0.95	1.08	1.05	0.95
σ_1	3.73	3.63	2.97	5.80	4.60	4.15	4.17
σ_2	1.11	2.32	1.59	1.47	1.28	2.60	2.20
σ_3	0.59	0.33	0.47	0.37	0.59	0.60	0.29
σ_4	0.007	0.051	0.050	0.042	0.021	0.023	0.052

807 Table 6 Same as Table 2, but for PL.

808

Types	Z_h (dBZ)	Z_{dr} (dB)	K_{dp} (deg km ⁻¹)	ρ_{hv}
PR	-10.4~0.4	2.8~5.8	-0.3~0.0	0.90~0.99
AG	-1.5~9.1	-0.5~3.9	-0.5~0.0	0.92~0.99
RD	-8.1~6.9	0.5~2.3	-0.2~0.2	0.95~0.99
IR	-10.8~5.8	-0.8~4.9	-0.6~0.0	0.96~0.99
PL	-16.4~-9.4	2.8~7.1	-0.2~-0.1	0.95~1.08

809 Table 7 Thresholds of polarimetric variables to classify PR, AG, RD, IR, and PL. The
810 estimates of ρ_{hv} larger than 1 are due to the signal processing technique.

811

List of Figures

Figure 1. Example photographs of identified ice particles: (a) pristine dendrites, (b) aggregates of dendrites, (c) rimed dendrites, (d) irregular crystals, and (d) plate-like crystals.	41
Figure 2 Flowchart of the clustering procedures.	42
Figure 3 Probability density functions for pristine dendrites (PR). The projection of the five-dimensional distribution is displayed: (a) Z_h - Z_{dr} , (b) Z_h - K_{dp} , and (c) Z_h - ρ_{hv} . The black dots represent the prototype means.	43
Figure 4 Same as Figure 3, but for aggregates of dendrites (AG).	44
Figure 5 Same as Figure 3, but for rimed dendrites (RD).	45
Figure 6 Same as Figure 3, but for irregular crystals (IR).	46
Figure 7 Same as Figure 3, but for plate-like crystals (PL).	47
Figure 8 Comparison between PDFs of pristine dendrites (PR), aggregates of dendrites (AG), rimed dendrites (RD), irregular crystals (IR), and plate-like crystals (PL).	48
Figure 9 Hemispheric Range Height Indicator (HRHIs) scans of (a) Z_h , (b) Z_{dr} , (c) K_{dp} , and (d) ρ_{hv} . Data were collected at a 52.2° azimuthal angle on 1611 UTC 02 May 2013.....	49
Figure 10 Comparison between results of (a) the Maximum Prototype Likelihood Classifier, (b) the simple Bayesian classifier, and (c) the Bayesian method based on Markov random field modeling. Data were collected at a 52.2° azimuthal angle on 1611 UTC 02 May 2013.	50
Figure 11 Same as Figure 9, but for HRHI data collected at a 52.2° azimuthal angle on 1836 UTC 06 May 2013.	51
Figure 12 Same as Figure 10, but for HRHI data collected at a 52.2° azimuthal angle on 1836 UTC 06 May 2013.	52
Figure 13 Same as Figure 9, but for PPI data collected at a 0.2° elevation angle on 1930 UTC 06 May 2013.	53
Figure 14 Same as Figure 10, but for PPI data collected at a 0.2° elevation angle on	

841 1930 UTC 06 May 2013.54

842

843

844 Figure 1. Example photographs of identified ice particles: (a) pristine dendrites, (b)
845 aggregates of dendrites, (c) rimed dendrites, (d) irregular crystals, and (e) plate-like
846 crystals.

847

848

849 Figure 2 Flowchart of the clustering procedures.

850

Figure 3 Probability density functions for pristine dendrites (PR). The projection of the five-dimensional distribution is displayed: (a) Z_h – Z_{dr} , (b) Z_h – K_{dp} , and (c) Z_h – ρ_{hv} . The black dots represent the prototype means.

Figure 4 Same as Figure 3, but for aggregates of dendrites (AG).

Figure 5 Same as Figure 3, but for rimed dendrites (RD).

Figure 6 Same as Figure 3, but for irregular crystals (IR).

Figure 7 Same as Figure 3, but for plate-like crystals (PL).

Figure 8 Comparison between PDFs of pristine dendrites (PR), aggregates of dendrites (AG), rimed dendrites (RD), irregular crystals (IR), and plate-like crystals (PL).

Figure 9 Hemispheric Range Height Indicator (HRHIs) scans of (a) Z_h , (b) Z_{dr} , (c) K_{dp} , and (d) ρ_{hv} . Data were collected at a 52.2° azimuthal angle on 1611 UTC 02 May 2013.

Figure 10 Comparison between results of (a) the Maximum Prototype Likelihood Classifier, (b) the simple Bayesian classifier, and (c) the Bayesian method based on Markov random field modeling. Data were collected at a 52.2° azimuthal angle on 1611 UTC 02 May 2013.

Figure 11 Same as Figure 9, but for HRHI data collected at a 52.2° azimuthal angle on 1836 UTC 06 May 2013.

Figure 12 Same as Figure 10, but for HRHI data collected at a 52.2° azimuthal angle on 1836 UTC 06 May 2013.

Figure 13 Same as Figure 9, but for PPI data collected at a 0.2° elevation angle on 1930 UTC 06 May 2013.

Figure 14 Same as Figure 10, but for PPI data collected at a 0.2° elevation angle on 1930 UTC 06 May 2013.

(a)

(b)

906

907 Figure_Apx 1 (a) First-order neighborhoods system and (b) second-order

908 neighborhoods system

909

Figure_Apx 2 Cliques q_1 , q_2 , and q_3 associated with a second-order neighborhood system: (a) A second-order neighborhood system, (b) single-site clique, (c) pair-site clique, and (d) triple-size clique.