



Title	Fitting Precipitation Particle Size-Velocity Data to Mixed Joint Probability Density Function with an Expectation Maximization Algorithm
Author(s)	Katsuyama, Yuta; Inatsu, Masaru
Citation	Journal of atmospheric and oceanic technology, 37(5), 911-925 <a href="https://doi.org/10.1175/JTECH-D-19-0150.1">https://doi.org/10.1175/JTECH-D-19-0150.1</a>
Issue Date	2020-05
Doc URL	<a href="http://hdl.handle.net/2115/79670">http://hdl.handle.net/2115/79670</a>
Type	article
File Information	J. Atmos. Ocean. Technol.37-5_911-925.pdf



[Instructions for use](#)

# Fitting Precipitation Particle Size–Velocity Data to Mixed Joint Probability Density Function with an Expectation Maximization Algorithm

YUTA KATSUYAMA<sup>a</sup>

*Graduate School of Science, Hokkaido University, Sapporo, Japan*

MASARU INATSU

*Faculty of Science, and Center for Natural Hazards Research, Hokkaido University, Sapporo, Japan*

(Manuscript received 3 September 2019, in final form 23 March 2020)

## ABSTRACT

This paper proposes an estimation method of joint size and terminal velocity distribution on the basis of sampling data of precipitation particles containing multiple types. Assuming that the velocity follows the normal distribution and the size follows the gamma distribution, the method searches a locally maximum logarithmic likelihood within a realistic parameter range using the expectation–maximization algorithm. Several test populations were prepared with a realistic number of elements, and then the method was evaluated by retrieving the populations from their sample. The results showed that the original parameters were successfully estimated in most cases of the test population containing some of liquids, graupels, and rimed and unrimed aggregates. The original number of elements was also estimated with an adjustment of the number of elements in a manner such that each of their minority fractions exceeded a threshold. Applied to the two-dimensional disdrometer observation data, the method was helpful to discard frequently observed erroneous data with unrealistically large fall velocity.

## 1. Introduction

The precipitation particle size distribution (PSD) was already well known for the liquid phase (Marshall and Palmer 1948) and for the solid phase (Gunn and Marshall 1958). The number concentration rate of precipitation particles diameter  $N_d$  ( $\text{m}^{-3} \text{mm}^{-1}$ ) follows a parametric gamma distribution on the diameter  $D$  (mm) as

$$N_d(D) = N_0 D^\mu \exp(-\lambda D), \quad (1)$$

where  $N_0$  ( $\text{m}^{-3} \text{mm}^{-1}$ ),  $\mu$ , and  $\lambda$  ( $\text{mm}^{-1}$ ) denote the intercept parameter, the distribution shape parameter, and the slope parameter, respectively. A long-lasting problem is how observed samplings of precipitation particle size are fit to a parametric distribution. Any observation instruments available at present tended to miss small-size particles for their insufficient sensitivity

and resolution, unless disdrometers like Meteorological Particle Sensor (Baumgardner et al. 2002) are combined to cover small sizes; this is hereinafter referred to as truncation problem (Handwerker and Straub 2011). Moreover large-size particles were rarely detected, which often brings us a kind of sampling problems. The moment method has been widely used for the fitting to size distribution (e.g., Vivekanandan et al. 2004; Smith and Kliche 2005; Yuter et al. 2006; Brandes et al. 2007; Kliche et al. 2008; Cao and Zhang 2009; Smith et al. 2009; Handwerker and Straub 2011), but a shortage of sampling length biases the estimate of parameters (Smith and Kliche 2005). The maximum likelihood and L-moment method can provide an unbiased estimate, but the truncation problem destabilizes the statistics (Kliche et al. 2008; Cao and Zhang 2009). Smith et al. (2009) showed that the above methods were difficult to fit limited sampling data to a gamma distribution of a small shape parameter. Recently, Yano et al. (2018) relieved this problem in the fitting to a gamma distribution solely on the basis of bulk quantities of PSD by applying the maximum entropy principle (Yano et al. 2016). Another problem comes up when fitting the mixture of multiple particle types in a single sampling (Brandes et al. 2007). For example, sampling data possibly contain

<sup>a</sup> Current affiliation: Tohkamachi Experimental Station, Forestry and Forest Products Research Institute, Tohkamachi, Japan.

Corresponding author: Yuta Katsuyama, ykatsuyama2020@affrc.go.jp

graupels and aggregates in a time interval (Yuter et al. 2006) that should be long enough to avoid the sampling problem (Smith and Kliche 2005).

On the other hand, the particle terminal velocity  $V$  ( $\text{m s}^{-1}$ ) is assumed to be proportional to the power of the particle's diameter as

$$V(D) = a_v D^{b_v}, \quad (2)$$

where  $a_v$  and  $b_v$  depend on particle phase, shape, and degree of riming (Atlas et al. 1973; Locatelli and Hobbs 1974). It is remarked that another functional form may be possible for liquid particles (Atlas et al. 1973) but we chose the simpler form as Eq. (2) to avoid a superfluous parameter throughout the paper. Many publications have been devoted to find an optimal fitting curve for particle diameter and velocity, on the basis of observed sampling data with optical or video disdrometers such as Particle Size Velocity (PARSIVEL; Löffler-Mang and Joss 2000; Battaglia et al. 2010; Angulo-Martínez et al. 2018), Laser Precipitation Monitor (LPM; Angulo-Martínez et al. 2018), and two-dimensional video disdrometer (2DVD; Kruger and Krajewski 2002). Recently, Molthan and Colle (2012) compared on-site observation with model outputs in every separated cell on the diameter–velocity quarter plane, instead of the direct comparison of velocity–diameter relationship or probability density function (PDF) on the plane. Bernauer et al. (2016) classified snow particles observed by 2DVD into three ranges of riming degree and four crystal types. The above fitting to a power law like Eq. (2) is, however, very difficult in a case of mixed particle types, because the least squares method conventionally used assumed a single type of precipitation particles. Enough of a short time interval, like 1 min (Bernauer et al. 2016), could partly, not perfectly, avoid the mixing, around the period when two independent distributions are clearly transited, by not sampling data from both the distributions, but would cause the sampling problem. A fitting to mass flux, a production of size–velocity and size–mass relations, partially resolved this problem (Ishizaka et al. 2013), but it is still difficult in a case of a clear bimodal distribution due to coexistence of liquid and solid phases (Brandes et al. 2007; Yuter et al. 2006) or coexistence of graupels and aggregates. Another problem in the fitting to size–velocity relation is a sampling error in small-size particles with an unrealistically high speed. This was often found in the output of single-sensor type disdrometers such as PARSIVEL and LPM (Angulo-Martínez et al. 2018), when particles passed through the edge of the sensitivity frame (Minda et al. 2016). This was also found in a double line-sensor type disdrometer of 2DVD for a failure of matching (Huang

et al. 2010; Bernauer et al. 2015). One inevitably excluded the part of sampling data out of the velocity range between 0.5 and 5  $\text{m s}^{-1}$  (Huang et al. 2010) and excluded data of small size less than 1 mm (Yuter et al. 2006; Bernauer et al. 2015, 2016), although the method filtering drops outside the  $\pm 40\%$  of the theoretical fall velocity may be also applicable if all particles are guaranteed to be liquid (Kruger and Krajewski 2002). This treatment deteriorated the truncation and sampling problems in the fitting to a size distribution, however.

This study aims to relieve the above problems of the mixing and the erroneous data often observed by disdrometers by fitting diameter–velocity data including multiple precipitation particle types and phases to mixed joint PDF on the diameter–velocity quarter plane; We conveniently utilized the gamma and normal distributions and the velocity–diameter relationship to evaluate the PDF in the comparison with a conventional method and well-known velocity–diameter relationships previously proposed. The fitting problem can be then reduced to a numerical search of an optimal set of parameters of the mixed joint PDF from the diameter–velocity data. We here develop the method with an aid of the expectation–maximization (EM) algorithm (Dempster et al. 1977). A performance of the new method is evaluated by a PDF retrieval from test data randomly sampled from a given population PDF. We next emphasize the strength of the new method to discern the coexistence of precipitation particle types and the sampling error in a subset of 2DVD data. This paper is organized as follows. Section 2 describes the observed 2DVD data, and section 3 elucidates the proposed method including a nutshell of the EM algorithm. We design experiments to check performance of the new method in section 4 and detect a bundle of erroneous data by applying the method to sample datasets obtained by the 2DVD in section 5. Section 6 discusses the limitation of the new method and section 7 concludes this paper.

## 2. Observation data

We used the precipitation diameter–velocity data obtained by a 2DVD installed at a site of the Institute of Low Temperature Science, Hokkaido University, Sapporo, Japan, located at 43.083°N, 141.339°E (Nagumo and Fujiyoshi 2015; Campbell et al. 2018). The 2DVD (Kruger and Krajewski 2002), widely used in the scientific community (Huang et al. 2010; Bernauer et al. 2015, 2016), measures precipitation particle diameter and velocity by matching two images captured by double line sensors installed at an upper and a lower position of the instrument. Image processing algorithm originally implemented in the 2DVD sensor matched the two particle images. As a preprocessing, we discarded erroneous

sampling particles with their velocity larger than  $15 \text{ m s}^{-1}$ , much larger than the terminal velocity of rain droplets (Atlas et al. 1973). We used surface air temperatures and precipitation records by Sapporo Meteorological Observatory of the Japan Meteorological Agency (JMA) located at  $43.0605^\circ\text{N}$ ,  $141.329^\circ\text{E}$ ,  $2.6 \text{ km}$  away from the 2DVD site (JMA 2019a,b).

### 3. Method

#### a. Mixed joint PDF

The PDF of particle diameter was here assumed to be the gamma distribution, a normalized version of Eq. (1):

$$P(D|\mu, \lambda) = \frac{\lambda^{\mu+1} D^\mu}{\Gamma(\mu+1)} \exp(-\lambda D), \quad (3)$$

where  $\Gamma$  is the gamma function. Here  $P(D|\mu, \lambda)$  indicates conditional probability density of  $D$  under parameters  $\mu$  and  $\lambda$ . We assume that the particle's terminal velocity follows the normal distribution with its mean being Eq. (2):

$$P(V|D, a_v, b_v, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(V - a_v D^{b_v})^2}{2\sigma^2}\right], \quad (4)$$

where  $\sigma^2$  is variance of velocity. This assumption was supported by theoretical experiments (Schmitt et al. 2019) and observations (Sasyo and Matsuo 1980). In multiplying Eqs. (3) and (4), a joint PDF of particle diameter and velocity is

$$P(V, D|a_v, b_v, \sigma^2, \mu, \lambda) = \frac{\lambda^{\mu+1} D^\mu}{\sqrt{2\pi\sigma^2} \Gamma(\mu+1)} \times \exp\left[-\frac{(V - a_v D^{b_v})^2}{2\sigma^2} - \lambda D\right]. \quad (5)$$

Linearly combining independent joint PDFs, a mixed joint PDF with  $K$  elements was written as

$$P(V, D|\theta) = \sum_{k=1}^K \omega_k P(V, D|a_{vk}, b_{vk}, \sigma_k^2, \mu_k, \lambda_k), \quad (6)$$

where  $\omega_k$  is a mixing fraction of the  $k$ th PDF element and should satisfy  $\sum_{k=1}^K \omega_k = 1$ . Now we introduced the vector  $\theta$  whose components are conditional parameters as

$$\theta = (\omega_1 \dots \omega_K a_{v1} \dots a_{vK} b_{v1} \dots b_{vK} \sigma_1^2 \dots \sigma_K^2 \mu_1 \dots \mu_K \lambda_1 \dots \lambda_K)^T.$$

Many possible PDFs can be constructed by sweeping a set of parameters  $\theta$  within a realistic range. Given a

sampling dataset, we may search an optimal parameter set  $\theta$  to estimate the PDF to be fit to the data. The distribution of particle number concentration rate within diameter–velocity quarter plane,  $N_{\text{vd}} (\text{m}^{-3} \text{s m}^{-1} \text{mm}^{-1})$ , more suitable unit for the traditional definition in meteorology, was easily obtained by multiplying the number concentration of total precipitation particles  $N_{\text{tot}} (\text{m}^{-3})$ :

$$N_{\text{vd}}(V, D) = N_{\text{tot}} P(V, D|\theta). \quad (7)$$

Note that the PDF [Eq. (6)] and the number concentration rate [Eq. (7)] explained diameter–velocity data by a disdrometer directly observing particle number concentration (e.g., Muramoto and Shiina 1989).

The joint PDF following diameter–velocity data by disdrometers observing precipitation particles passing a finite area such as 2DVD, PARSIVEL, and LPM is described as

$$P_S(V, D|a_v, b_v, \sigma^2, \mu, \lambda) = \frac{V}{a_v D^{b_v}} P(V, D|a_v, b_v, \sigma^2, \mu + b_v, \lambda). \quad (8)$$

From this equation, the shape parameter  $\mu$  is shifted by  $b_v$  with a diameter–velocity relationship form of Eq. (2) (Ignaccolo and De Michele 2014; Adirosi et al. 2016) and  $P_S(V, D)$  extends to the tendency of higher velocity due to a factor of  $V/a_v D^{b_v}$ . The mean fall velocity for the joint PDF of Eq. (8),  $V_S$ , is also different from Eq. (2):

$$V_S(D) = \frac{\int_{-\infty}^{\infty} V P_S(V, D|a_v, b_v, \sigma^2, \mu, \lambda) dV}{\int_{-\infty}^{\infty} P_S(V, D|a_v, b_v, \sigma^2, \mu, \lambda) dV} = \frac{\sigma^2}{a_v D^{b_v}} + a_v D^{b_v}. \quad (9)$$

The mixed joint PDF based on Eq. (8) is

$$P_S(V, D|\theta) = \sum_{k=1}^K \omega_k P_S(V, D|a_{vk}, b_{vk}, \sigma_k^2, \mu_k, \lambda_k). \quad (10)$$

The distribution of particle number concentration rate is

$$N_{\text{vd}}(V, D) = \frac{L}{A \Delta t V} P_S(V, D|\theta), \quad (11)$$

where  $A$ ,  $\Delta t$ , and  $L$  are, respectively, an observing area ( $\text{m}^2$ ), the sampling time (s), and the number of total particles observed by a disdrometer. See appendix A for derivations of Eqs. (8) and (11).

Equations (5) and (8) should be properly assumed as a joint PDF form that diameter–velocity data by a

disdrometer should follow. Because we could not formulate the EM algorithm for parameters  $a_v$  and  $b_v$  with this joint PDF form of Eq. (8) (cf. section 3b), however, the PDF form of Eq. (5) was applied to the observed sampling data by 2DVD. The problem on using Eq. (5) for 2DVD data will be discussed in section 6.

*b. EM algorithm*

The EM algorithm (Dempster et al. 1977) was used to search an optimal parameter set of Eq. (6). First, we posed a kernel PDF from the  $l$ th particle for the  $k$ th PDF element as  $P(V_l, D_l|a_{vk}, b_{vk}, \sigma_k^2, \mu_k, \lambda_k)$ . The logarithmic (log) likelihood was then constructed by summing the logarithm of this kernel PDFs for all of  $L$  particles in the sample:

$$\ln\Phi(\mathbf{V}, \mathbf{D}|\boldsymbol{\theta}) = \sum_{l=1}^L \ln \left[ \sum_{k=1}^K \omega_k P(V_l, D_l|a_{vk}, b_{vk}, \sigma_k^2, \mu_k, \lambda_k) \right]. \tag{12}$$

The EM algorithm finds a parameter set of  $\boldsymbol{\theta}$  that satisfies

$$\frac{\partial}{\partial \boldsymbol{\theta}} \ln\Phi(\mathbf{V}, \mathbf{D}|\boldsymbol{\theta}) = 0 \tag{13}$$

to maximize the log likelihood [Eq. (12)] under the constraint of  $\sum_{k=1}^K \omega_k = 1$ . Equation (13) is reduced to the iteration relations below:

$$a_{vk} = \frac{\sum_{l=1}^L \gamma_{lk} V_l D_l^{b_{vk}}}{\sum_{l=1}^L \gamma_{lk} D_l^{2b_{vk}}} \tag{14}$$

$$\frac{\partial}{\partial b_{vk}} \ln\Phi(\mathbf{V}, \mathbf{D}|\boldsymbol{\theta}) = \frac{a_{vk}}{\sigma_k^2} \sum_{l=1}^L \gamma_{lk} \ln(D_l) (V_l - a_{vk} D_l^{b_{vk}}) D_l^{b_{vk}} = 0 \tag{15}$$

$$\sigma_k^2 = \frac{\sum_{l=1}^L \gamma_{lk} (V_l - a_{vk} D_l^{b_{vk}})^2}{\sum_{l=1}^L \gamma_{lk}}, \tag{16}$$

$$\ln(\mu_k + 1) - \Psi(\mu_k + 1) = \ln(\overline{D}_k) - \overline{\ln(D)}_k, \tag{17}$$

$$\lambda_k = \frac{\mu_k + 1}{\overline{D}_k}, \text{ and} \tag{18}$$

$$\omega_k = \frac{\sum_{l=1}^L \gamma_{lk}}{L}, \tag{19}$$

where

$$\Psi(x) = \frac{d}{dx} \ln[\Gamma(x)]$$

is the digamma function, the responsibility fraction  $\gamma_{lk}$  ranging over 0–1 as

$$\gamma_{lk} = \frac{\omega_k P(V_l, D_l|a_{vk}, b_{vk}, \sigma_k^2, \mu_k, \lambda_k)}{\sum_{j=1}^K \omega_j P(V_l, D_l|a_{vj}, b_{vj}, \sigma_j^2, \mu_j, \lambda_j)} \tag{20}$$

represents how much the  $l$ th particle explains the  $k$ th PDF element, and  $\overline{D}_k$  and  $\ln(D)_k$ , respectively, represent the mean diameter  $[ = (\sum_{l=1}^L \gamma_{lk} D_l) / (\sum_{l=1}^L \gamma_{lk}) ]$  and the mean log diameter  $\{ = [\sum_{l=1}^L \gamma_{lk} \ln(D_l)] / (\sum_{l=1}^L \gamma_{lk}) \}$  corresponding to the  $k$ th PDF element. We numerically solved Eq. (15) by the secant method (Householder 1970), and solved Eq. (17) by iteration (Kliche et al. 2008):

$$\mu_k + 1 = (\mu_k + 1) \frac{\ln(\mu_k + 1) - \Psi(\mu_k + 1)}{\ln(\overline{D}_k) - \overline{\ln(D)}_k}. \tag{21}$$

Note that Eqs. (17) and (18) with  $K = 1$  are the completely same form as the equations of the conventional maximum-likelihood method in estimating particle size distribution (Kliche et al. 2008) so that the EM algorithm with  $K = 1$  estimates  $\mu$  and  $\lambda$  as comparably as the maximum-likelihood method.

Next, the EM algorithm estimates an optimal parameter set including responsibility fraction [Eq. (20)] in the iteration procedure. Initially, we give the number of PDF elements  $K$  and parameters for the  $k$ th PDF element as

$$(\ln a_{vk}, b_{vk}, \sigma_k^2, \mu_k, \lambda_k, \omega_k) = \begin{cases} \left[ \ln(\check{a}_v) + \ln\left(\frac{\hat{a}_v}{\check{a}_v}\right) \frac{k-1}{K-1}, \frac{(k-1)}{K-1}, V', \frac{\overline{D}^2}{D'} - 1, \frac{\overline{D}}{D'}, \frac{1}{K} \right] & (K > 1) \\ \left[ \frac{\ln(\hat{a}_v \check{a}_v)}{2}, 0.5, V', \frac{\overline{D}^2}{D'} - 1, \frac{\overline{D}}{D'}, 1 \right] & (K = 1) \end{cases}, \tag{22}$$

where

$$\begin{cases} \hat{a}_v = \frac{\max_{1 \leq l \leq L} V_l}{D} \\ \check{a}_v = \frac{\bar{V}}{4} \end{cases}, \quad (23)$$

with overbar and prime denoting the average and the variance for all the particles. See [appendix B](#) for details of this initial parameter determination. We call providing the responsibility fraction for the parameter set with Eq. (20) the E step and updating the parameter set with a sequential calculation order of  $b_v$  with Eq. (15) substituting Eq. (14) for  $a_v$ ,  $a_v$  with Eq. (14),  $\sigma^2$  with Eq. (16),  $\mu$  with Eq. (17),  $\lambda$  with Eq. (18), and  $\omega$  with Eq. (19) the M step. The E step and M step are iterated if the log likelihood increases by more than  $10^{-2}$  relative to the value before the steps. Otherwise the iteration stops. No more iterations than 5000 avoids computation for divergent cases. It is noted that a convergence solution by this iteration procedure is solely a locally most likelihood nearest from the initial and is not always the most likelihood globally, and therefore the EM algorithm is not referred to as the maximum-likelihood method.

#### c. Estimation of the number of PDF elements

A remaining parameter that should be estimated is the number of PDF elements,  $K$ , even though the EM algorithm itself can be performed with given  $K$ . We now introduce the Akaike information criterion (AIC; [Akaike 1974](#))

$$\text{AIC} = \ln \Phi(\mathbf{V}, \mathbf{D} | \boldsymbol{\theta}) - 6K \quad (24)$$

and Bayesian information criterion (BIC; [Schwarz 1978](#))

$$\text{BIC} = \ln \Phi(\mathbf{V}, \mathbf{D} | \boldsymbol{\theta}) - 3K \ln L. \quad (25)$$

An appropriate number of PDF elements can be searched by maximizing AIC or BIC by sweeping  $K$  from 1 to 4. We also applied minority rejection method to find the appropriate number of elements by an adjustment such that the smallest mixing fraction in the elements is larger than a threshold following 0.1  $K$ . See [appendix C](#) for the rationality of this choice.

## 4. Performance test

### a. Design

Performance of the proposed method is examined in light of (i) computational stability of the EM algorithm, (ii) the accuracy on the parameter forming the PDF elements, (iii) the accuracy on the estimate of the optimal

TABLE 1. Parameters of the variance  $\sigma^2$  and the slope  $\lambda$  for populations with small variance and small slope (label 0), small variance and large slope (label s), and large variance and small slope (label v) for the probability density function elements of liquids, graupel, rimed aggregates, and unrimed aggregates (labels L, G, R, and U, respectively).

Population type	0		s		v	
	$\sigma^2$	$\lambda$	$\sigma^2$	$\lambda$	$\sigma^2$	$\lambda$
L	0.16	4	0.16	6.5	0.64	4
G	0.04	1.5	0.04	2.5	0.16	1.5
R	0.04	0.4	0.04	0.8	0.16	0.4
U	0.04	0.4	0.04	0.8	0.16	0.4

PDF-element number, and (iv) comparison with the least squares method conventionally used. All points can be checked by experiments, similarly designed, to estimate a given PDF from random sampling data from the PDF, just like a Monte Carlo simulation ([Smith and Kliche 2005](#); [Kliche et al. 2008](#); [Cao and Zhang 2009](#); [Smith et al. 2009](#); [Handwerker and Straub 2011](#)).

For points i–iii, we prepared nine population distributions containing PDF elements including liquid precipitation particle L with  $(a_{vL}, b_{vL}) = (3.78, 0.67)$  ([Atlas et al. 1973](#)), lump graupel G with  $(a_{vG}, b_{vG}) = (1.3, 0.66)$  ([Locatelli and Hobbs 1974](#)), densely rimed aggregates R with  $(a_{vR}, b_{vR}) = (1.1, 0.15)$  ([Ishizaka 1995](#)), or side-plane unrimed aggregates U  $(a_{vU}, b_{vU}) = (0.82, 0.12)$  ([Locatelli and Hobbs 1974](#)). The mixed joint PDFs prepared here are mixture of liquid particle, graupel, and rimed aggregates (LGR), mixture of graupel and rimed aggregates (GR), and mixture of rimed and unrimed aggregates (RU), with the same mixing fraction for all elements (i.e.,  $\omega = 1/3$  for LGR and  $\omega = 1/2$  for GR and RU). Note that these combinations of particle types were subjectively chosen to represent some of possible observed examples. For point iv, a single PDF element with liquids (L), graupels (G), rimed aggregates (R), or unrimed aggregates (U) was prepared. We set the shape parameter as  $\mu = 1.5$  for all of the populations. Rather than the parameters  $(a_v, b_v, \omega, \mu)$ , the PDF elements depend on data variance  $\sigma^2$  and slope  $\lambda$ . We set three different combinations of these two parameters to check a variety of mixed joint PDFs as [Table 1](#). Three kinds of PDF element combined with three parameter sets of  $\sigma^2$ ,  $\mu$ , and  $\lambda$  made nine mixed joint PDFs ([Fig. 1](#)), and each experiment is nominated as the prefix of LGR, GR, or RU combined with the suffix of 0, s, or v. For example, the experiment LGRv means a given PDF of elements LGR with a parameter set of v. Next, we randomly sampled diameter–velocity data from the population distribution by von Neumann’s method within ranges of  $0 \leq D \leq 30$  mm and  $0 \leq V \leq 10$  m s<sup>-1</sup>. The sensitivity to the sample size was checked by

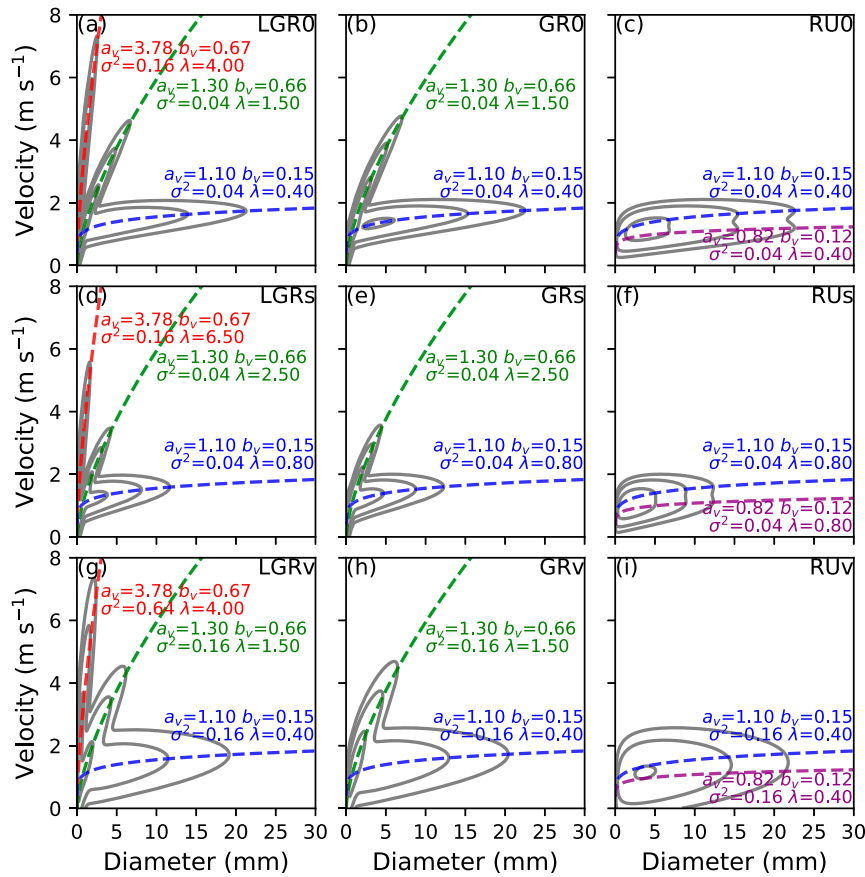


FIG. 1. Population distribution given in the performance tests of (a) LGR0, (b) GR0, (c) RU0, (d) LGRs, (e) GRs, (f) RUs, (g) LGRv, (h) GRv, and (i) RUV. Parameters of each PDF element are indicated by the text inset; diameter-velocity curves are drawn by dashed lines for liquids (red), graupel (green), rimed aggregates (blue), and unrimed aggregates (purple). Contours indicate probability of 0.1, 0.01, and 0.001. See the text for the details of the given population PDFs.

taking 100, 1000, 2000, 3000, 6000, and 10 000 particles. Last, the EM algorithm was applied to the sampled data and the obtained PDF was compared with the original. Each estimation was repeated 1000 times with different sampling data. To keep the same element order between estimate and population, the estimated elements are matched to the elements in the population by selecting the combination of  $(j, k)$  with the nearest neighbor measured by normalized Euclidian distance of

$$\left\| \left[ \begin{array}{c} (a_{vj} - a_{vk})/a_{vk} \\ (b_{vj} - b_{vk})/b_{vk} \end{array} \right] \right\|. \quad (26)$$

Here  $j$  represents the  $j$ th PDF elements in the estimation ( $j = 1, 2, 3$  in LGR and  $j = 1, 2$  in GR and RU), and  $k$  represents PDF element L, G, R, or U.

Point i is to check computational stability measured by the convergence rate of EM algorithm, in which we counted how many trials were numerically converged in 1000 trials with a different sample dataset. This check should be required because the EM algorithm, searching a local optimal solution nearest from an initial value, not a *global* optimal solution, is not always converged. We regarded that the trial was successful when the computation was converged for every number of PDF elements given from 1 to 4. Point ii was to check the accuracy of the parameters forming the PDF elements. This was evaluated as the root-mean-square error (RMSE) of parameters,  $a_v$ ,  $b_v$ ,  $\sigma^2$ ,  $\mu$ ,  $\lambda$ , and  $\omega$ , of the estimated PDF elements from the reference of the population PDF given (Fig. 1). The relative error of log likelihood was also used as the measure. Point iii was to check the accuracy of the estimate of the number of PDF elements  $K$  from 1 to 4. When the

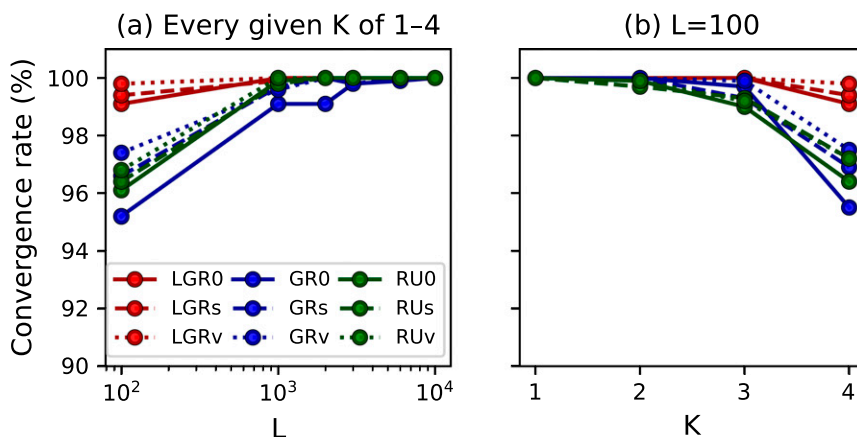


FIG. 2. The convergence rate (a) as a function of sample size  $L$  for every number of PDF elements  $K$  from 1 to 4 and (b) as a function of  $K$  in experiments with  $L = 100$ . The line color and type are as per the legend in (a).

number of PDF elements in the estimation was matched with the given number, we regarded it as hit. The hitting rate is defined as the hit number divided by the total trials that are converged. Point iv was to compare our method that could generally estimate a mixed joint PDF with the conventional method that estimated only a single velocity–diameter relationship. We then used the single-element PDF as the population with three different combination patterns of variance and slope parameters.

### b. Results

The computational stability was first checked by the convergence rate of estimation. Whatever we gave the PDF element number, the EM algorithm was converged by mostly 100% of 1000 trials if the sample size was greater than 1000 (Fig. 2a). In a small sample size of 100, the EM algorithm sometimes did not find a convergence solution with greater PDF elements given (Fig. 2b). Therefore, it is not recommended to give the number of PDF elements greater than 2 when the sample size is less than 100. Hereinafter we show the results with the sample data size at 2000.

Second, the accuracy of the estimation of PDF elements in comparison with the given population was checked by RMSE of each parameter (Fig. 3), in the estimation with the truth number of PDF elements  $K$  given. Greater the sample size was, more accurate the estimate was in most cases; the RMSE of any parameters plummeted with  $L$  exceeding 2000 (not shown). Parameters  $a_v$  and  $b_v$ , directly determined a diameter–velocity relationship, were successfully estimated, especially in the retrieval of graupel and rimed aggregates with small variance (GR0 and GRs; Figs. 3a,b). The

population containing rimed and unrimed aggregates with larger variance (RUV) was estimated with relatively larger RMSE. The variance parameter,  $\sigma^2$ , of PDF elements was also well retrieved with its RMSE almost less than a quarter of variance parameter in the PDF elements in the given population except for liquids for LGRv (Fig. 3c). The parameter  $\mu$  was estimated with its RMSE mostly less than 0.5 (Fig. 3d). The slope parameter  $\lambda$  was well retrieved in all the cases with its RMSE less than 10% of the slope parameter in the PDF elements in the given population (Fig. 3e). This estimate from smaller sample size than 2000 was much worse in RU0, RUs, and RUV (not shown). The RMSE of mixing fraction  $\omega$  ranged over 0.01–0.05 except for the estimation for RUV population (Fig. 3f).

Figure 4 shows the PDF of the relative error of log likelihood based on 1000 trials, in which positive and negative values mean overfitting and mismatching, respectively. None of the trials fell into a wrong optimum except for RU0 and RUs. In the cases of RU0 and RUs, the distribution of relative error shows bimodality separated at zero (Figs. 4c,f). If the relative error of log likelihood below  $-0.2\%$  was regarded as a wrong optimum, about 10% of the trials failed to estimate the true PDF. Because the error majority ranged from 0% to 0.5%, the PDF estimate was slightly biased to the overfitting side. It is worthwhile noting that this width of range was probably attributed to the uncertainty due to a sampling randomness indwelling in the sampled data, which was small enough here (Fig. 4). Therefore, the overfitting was the primary cause for the RMSE in estimated parameters (Fig. 3), but the mismatching was the secondary cause only in the cases of RU0 and RUs.

Third, we evaluated the number of PDF element by the hitting rate (Fig. 5). The result showed that the minority rejection method based on smallest mixing



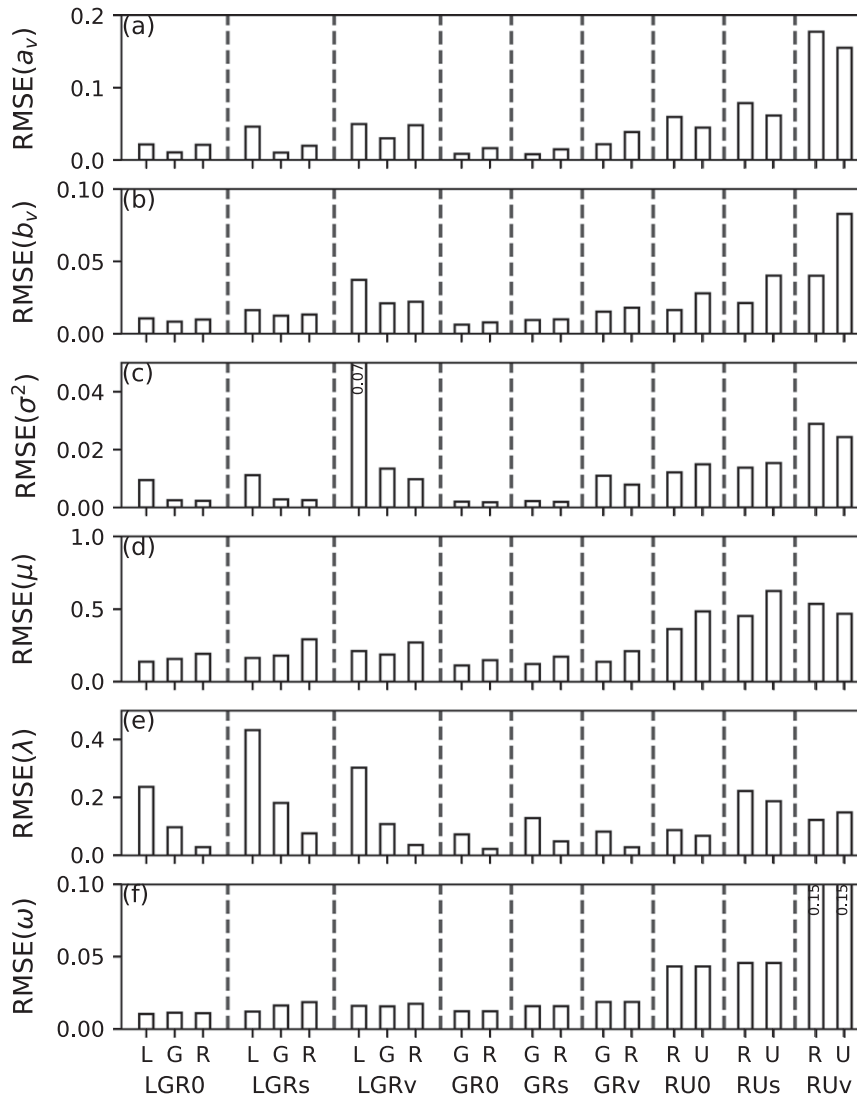


FIG. 3. RMSE of parameters (a)  $a_v$ , (b)  $b_v$ , (c)  $\sigma^2$ , (d)  $\mu$ , (e)  $\lambda$ , and (f)  $\omega$  in the estimate of nine populations for PDF elements of liquids (L), graupel (G), rimed aggregates (R), and unrimed aggregates (U), with the sampling size at 2000.

fraction provided the best performance in all the population patterns given (Fig. 5c). The hitting rate was asymptotic to 100% for greater sample size. The sample size of 1000 or more made the hitting rate of 100% for the populations of LGR0, LGRs, LGRv, GR0, GRs, and GRv based on the method maximizing BIC (Fig. 5b). The method maximizing AIC provided the worst estimates (Fig. 5a), very sensitive to the sample size; the hitting rate is too low for the population with the large variance parameter even the sample size is large.

Fourth, our method for the estimate from the population containing the single PDF element was compared with the conventional method to find an optimal

velocity–diameter relationship by the least squares method: the comparison was done for two parameters  $a_v$  and  $b_v$  (Fig. 6). The populations regardless of parameters  $\sigma^2$  and  $\lambda$  were retrieved with the comparable RMSE between the methods. The difference between methods was insensitive to the sample size (not shown). Hence, our method can be also applicable to a single-element population in addition to a multiple-elements population. Comparing the result of single PDF element cases (Fig. 6) and multiple PDF elements cases (Fig. 3), the estimations for the joint PDF population of GR0 and GRs had as a good performance as two individual estimations for the single PDF population G0 or R0 (Figs. 3 and 6). In contrast, the estimation for the RUs and RUv populations

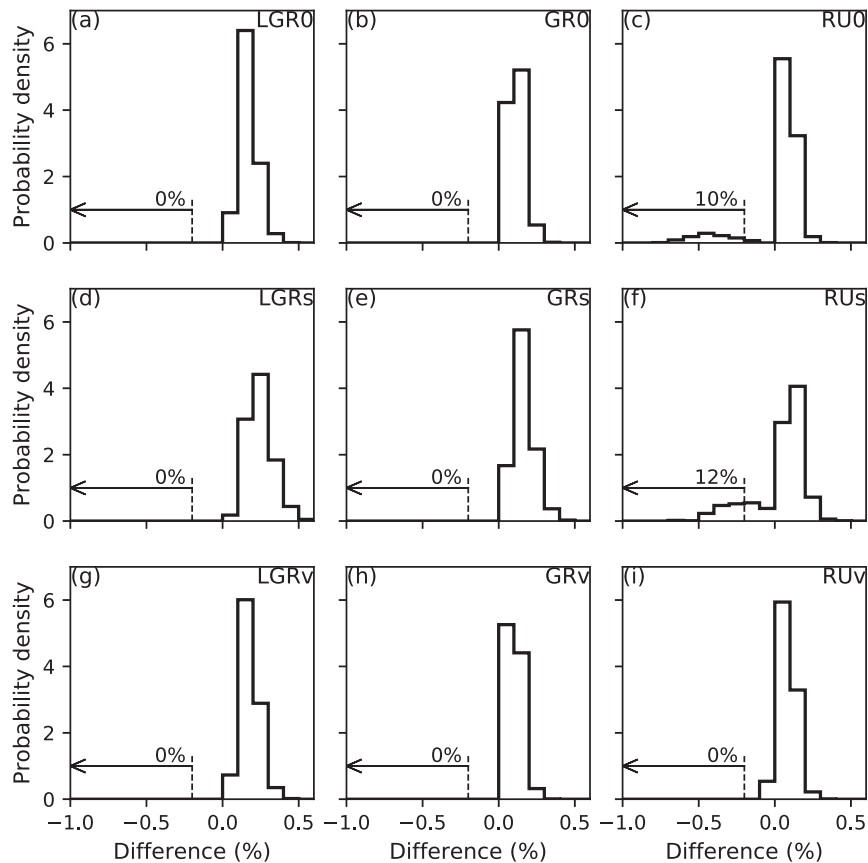


FIG. 4. The PDF of relative error of log likelihood for nine populations. Positive values mean overfitting, and negative values mean mismatching. Total probability integrated down from  $-0.2\%$  is indicated in the text inset.

had a worse performance than two estimations for Rs, Rv, Us, or Uv population.

## 5. Results

In addition to the performance test above, we here demonstrated how the proposed method fit to diameter–velocity data observed by the 2DVD at Sapporo (section 2), with two cases not to be considered in the performance test. The first case included erroneous and normal data during 1400–1430 Japan standard time (JST) 22 January 2017, when the JMA reported the snow particles fell in the air temperature of  $-4.9^{\circ}\text{C}$  (JMA 2019a). The second case included liquid and wet snow particles during 0100–0130 JST 6 December 2016, when the JMA reported that the rain turned to sleet and the air temperature changed from  $3^{\circ}$  to  $2.2^{\circ}\text{C}$  (JMA 2019b).

The observed data by the 2DVD during 1400–1430 JST 22 January 2017, included erroneous data with unrealistically large velocity probably due to the matching problem (Fig. 7a). With two PDF elements given, our method estimated one element (red in Fig. 7a) with its center at

$V = 2.61D^{0.54}$  and the variance of  $6.43\text{ m}^2\text{ s}^{-2}$ , which was so unrealistic. Moreover, this erroneous PDF element underestimated probability around  $1.5\text{ m s}^{-1}$  (red in Fig. 7c) due to the assumption that the velocity followed the normal distribution in this study. In contrast, the erroneous PDF well represented diameter distribution (red in Fig. 7b), because the matching error affected the velocity only. The estimation of the other PDF element (blue in Fig. 7a) was successfully fit to  $V = 0.83D^{0.21}$ , near to the optimal curve for graupel-like snow of hexagonal type,  $V = 0.86D^{0.25}$ , or aggregates of densely rimed radiating assemblages of dendrites,  $V = 0.79D^{0.27}$ , in Locatelli and Hobbs (1974) (Fig. 7a). This PDF element well represented the observed distributions of both the diameter and velocity (blue in Figs. 7b,c). However, because the estimated velocity–diameter relationship roughly diagnosed the characteristics of precipitation particles (Bernauer et al. 2016), the estimated velocity–diameter relationship  $V = 0.83D^{0.21}$  did not always guarantee the precipitation particles containing graupel-like snow and aggregates of densely rimed radiating assemblages. While the fitting result was

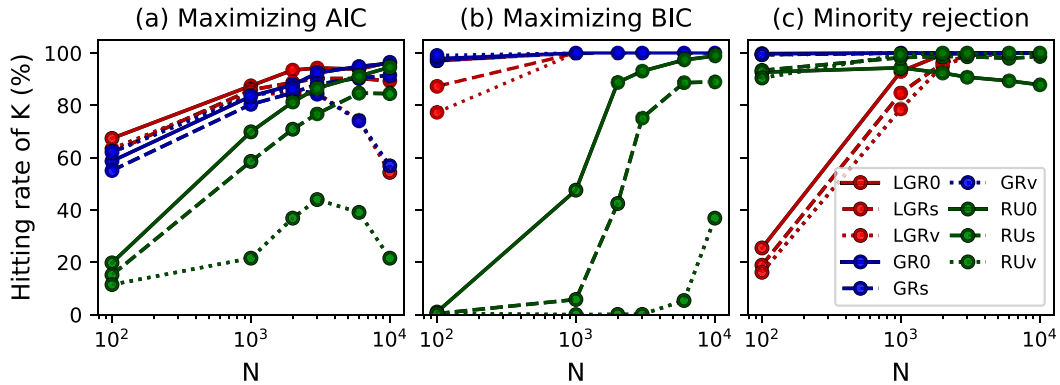


FIG. 5. Hitting rate of  $K$  estimated by (a) maximizing the AIC, (b) maximizing the BIC, and (c) the minority rejection method.

obtained with  $K = 2$  subjectively given, the minority rejection method estimated the number of PDF elements at  $K = 1$  because the minimum mixing fraction with  $K = 2$ ,  $\omega = 0.18$  (Fig. 7a), was less than the threshold of 0.2 determined in advance (section 3c).

Interesting is a different case for mixing rain droplets and wet snow particles observed with the 2DVD at Sapporo during 0100–0130 JST 6 December 2016 (Fig. 8). Here the minority rejection method estimated two PDF elements. If we subjectively fixed the number of elements at 3, we estimated the first element with its center at  $V = 2.33D^{0.03}$  and the variance of  $\sigma^2 = 0.13$  (green in Fig. 8), the second element with its center at  $V = 2.9D^{0.53}$  and the variance of  $\sigma^2 = 0.12$  (blue in Fig. 8), and the third element with its center at  $V = 2.83D^{0.62}$  and the variance of  $\sigma^2 = 3.46$  (red in Fig. 8). The first element resembles a distribution by wet snow particles reported by Yuter et al. (2006), and the second element was similar to a diameter–velocity curve of rain droplets ( $V = 3.78D^{0.67}$  in Atlas et al. 1973) with small variance and good fitness of velocity distribution (blue in Fig. 8c). The third element probably contained erroneous data because of its large variance and inappropriate fitting to velocity distribution (red in Fig. 8c), however. The minority rejection method estimated the number of elements at  $K = 2$ , probably because the number of erroneous data was smaller as shown in Fig. 7.

6. Discussion

We showed that the proposed method successfully reproduced the original PDF from the sampling data of its size larger than 2000 under a limitation of possible PDF element number considered (Fig. 3). However, as in the application to observation data including error (section 5), we must take a treatment

for erroneous data as the preprocessing because an arithmetic underflow often occurred in representing their probability. For example, the probability of particles at  $(D, V) = (2, 15)$  under a condition of  $(a_v, b_v, \sigma^2, \mu, \lambda) = (0.82, 0.12, 0.09, 1.5, 0.3)$  in Eq. (5) approximates  $10^{-481}$  that causes a double-precision underflow. This might cause the zero division error in the calculation of the responsibility fraction [Eq. (20)] with uniform mixing fractions among the PDF elements. This underflow problem can be relieved by setting the initial variance parameter larger than  $\sigma^2 = 0.16$  and by discarding particles with their velocity faster than  $15 \text{ m s}^{-1}$ . If one successfully avoided this problem by the preprocessing like above, the EM algorithm would make light of erroneous extreme data and search an optimal set of PDF elements that follow data majority.

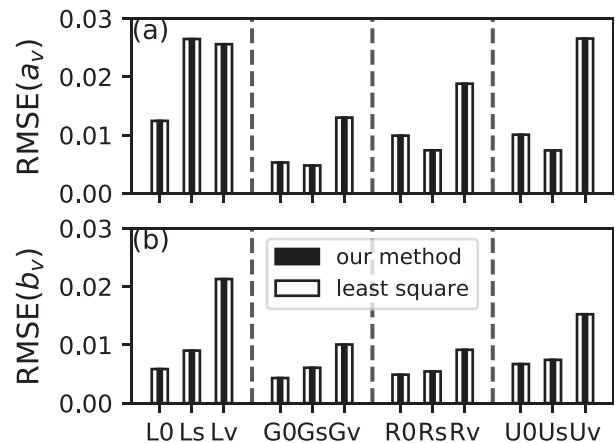


FIG. 6. RMSE of parameters (a)  $a_v$  and (b)  $b_v$  in the estimate of populations L0–Uv by our method (open bars) and the least squares method (filled bars).

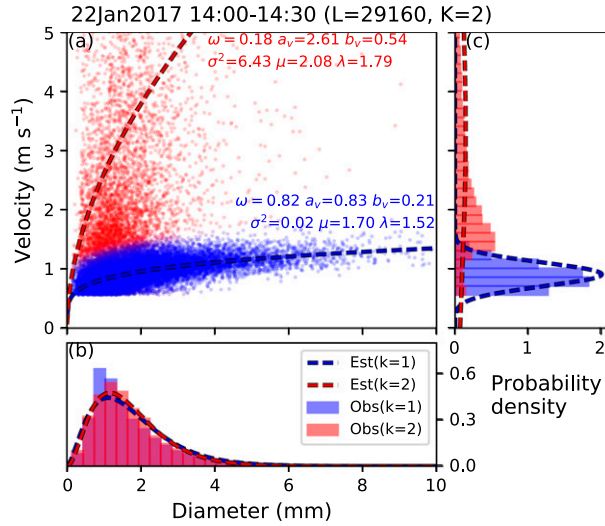


FIG. 7. (a) Scatterplot of diameter and velocity for precipitation particles observed by a 2DVD installed at Hokkaido University during 1400–1430 JST 22 Jan 2017. The colored dots show the responsibility fractions  $\gamma_{ik}$  for the two elements with blue and red and their mixture:  $(\gamma_{11}, \gamma_{12}) = (1, 0)$  with blue,  $(\gamma_{11}, \gamma_{12}) = (0, 1)$  with red, and  $(\gamma_{11}, \gamma_{12}) = (0.5, 0.5)$  with magenta. The text inset indicates the parameters for the PDF estimate with the number of PDF elements given at 2. The colored dotted lines are diameter–velocity curves estimated by our method. Also shown are the estimated (lines) and observed (bars) PDFs for (b) diameter and (c) fall velocity.

Moreover, we still suffer from the truncation effect attributed to not only the small size particles but also slow velocity ones. To relieve the problem, we now consider the truncated version of mixed joint PDF. The truncated version of particle size distribution is

$$P_T(D|\mu, \lambda) = \frac{P(D|\mu, \lambda)\Pi(D, D_{\min}, D_{\max})}{Q(D_{\max}|\mu, \lambda) - Q(D_{\min}|\mu, \lambda)}, \quad (27)$$

where  $D_{\min}$  and  $D_{\max}$  are, respectively, the minimum and maximum diameters that a disdrometer is capable to measure,  $\Pi$  is the rectangle function (1 between  $D_{\min}$  and  $D_{\max}$  and 0 elsewhere), and  $Q$  is the cumulative density function (Mallet and Barthes 2009; Johnson et al. 2014). Similar to Eq. (27), the truncated version of terminal velocity distribution is

$$P_T(V|D, a_v, b_v, \sigma^2) = \frac{P(V|D, a_v, b_v, \sigma^2)\Pi(V, V_{\min}, V_{\max})}{Q(V_{\max}|D, a_v, b_v, \sigma^2) - Q(V_{\min}|D, a_v, b_v, \sigma^2)}, \quad (28)$$

where  $V_{\min}$  and  $V_{\max}$  are, respectively, the minimum and maximum velocity that a disdrometer is capable to measure. Therefore, the truncated version of mixed

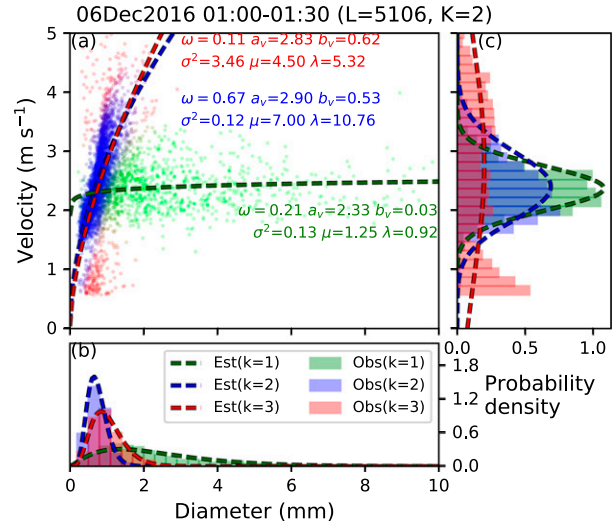


FIG. 8. (a) Scatterplot of diameter and velocity for precipitation particles observed by 2DVD during 0100–0130 JST 6 Dec 2016. The colored dots show the responsibility fractions  $\gamma_{ik}$  for the three elements with green, blue, and red and their mixture:  $(\gamma_{11}, \gamma_{12}, \gamma_{13}) = (1, 0, 0)$  with green,  $(\gamma_{11}, \gamma_{12}, \gamma_{13}) = (0, 1, 0)$  with blue,  $(\gamma_{11}, \gamma_{12}, \gamma_{13}) = (0, 0, 1)$  with red, and  $(\gamma_{11}, \gamma_{12}, \gamma_{13}) = (1/3, 1/3, 1/3)$  with gray. The text inset indicates the parameters for the PDF estimate with the number of PDF elements given at 3. The colored dotted lines are diameter–velocity curves estimated by our method. Also shown are the estimated (lines) and observed (bars) PDFs for (b) diameter and (c) fall velocity.

joint PDF is derived by linearly combing the joint PDF of  $P_T(V|D, a_v, b_v, \sigma^2) \times P_T(D|\mu, \lambda)$ . Modifying the EM algorithm using this truncated version of mixed joint PDF, the proposed method might relieve the truncation problem, but this is beyond the scope of this paper.

We next discuss the time interval for data sampling in the observation. The time interval that the total number of precipitation particles passing through an area of  $A$  ( $\text{m}^2$ ) reaches  $L$  is computed in appendix D as

$$\Delta t = \frac{L}{AR} \left[ \sum_{k=1}^K \frac{\omega_k \lambda^{b_{mk}} \Gamma(\mu_k + b_{vk} + 1)}{a_{mk} \Gamma(\mu_k + b_{vk} + b_{mk} + 1)} \right]^{-1}, \quad (29)$$

where  $R$  ( $\text{mg m}^{-2} \text{s}^{-1}$ ) is precipitation intensity,  $a_m$  and  $b_m$  are parameters of the mass–diameter relationship:  $M = a_m D^{b_m}$  (mg). Note that  $(a_m, b_m)$  for liquids, graupels and densely rimed and unrimed aggregates are (0.52, 3), (0.078, 2.8), (0.094, 1.9), and (0.04, 1.4), respectively (Atlas et al. 1973; Locatelli and Hobbs 1974; Ishizaka 1995). Figure 9 shows the time interval to obtain 2000 particles per a unit area of  $\sim 0.01 \text{ m}^2$  composed by the typical 2DVD (Kruger and Krajewski 2002). Since the sample size should be greater than 2000 for computational stability of the algorithm (Fig. 2) and

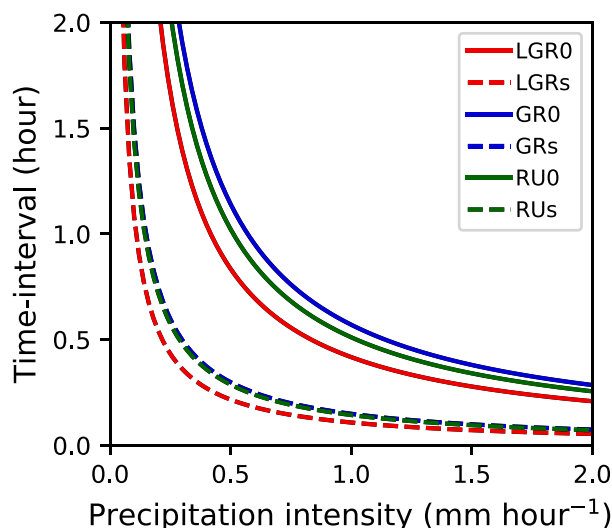


FIG. 9. Time interval (h) to obtain 2000 particles per a unit area ( $0.01 \text{ m}^2$ ) as a function of precipitation intensity ( $\text{mm h}^{-1}$ ).

for accuracy of the parameter estimation (Figs. 3 and 5), in  $1 \text{ mm h}^{-1}$  precipitation intensity, the time interval is roughly 5 min for LGRs, GRs, and RUs populations, 30 min for LGR0, LGRv, GR0, GRv, RU0, and RUv. Moreover, with an intake of PARSIVEL and LPM of almost one-half of 2DVD, the time interval is doubled (Angulo-Martínez et al. 2018).

We similarly compared our method with the conventional maximum-likelihood method by Kliche et al. (2008) in light of the shape  $\mu$  and slope  $\lambda$  parameters from the single population of liquid precipitation particles following the gamma distribution. They reported that the RMSE of shape parameter was 0.13 with respect to the population of  $\mu = 2$  for the sample size of 1000. This result was consistent with our result of the estimated PDF element corresponding to the liquid particles for the LGR0, LGRs, and LGRv populations (Fig. 3d). As for the slope parameter, Kliche et al. (2008) computed a median value of estimated  $\lambda$  normalized with its population being  $\sim 1.0$  with a sample size of 1000. In this study, the normalized  $\lambda$  of the estimated PDF element corresponding to the liquid particles was  $\sim 1.02$  for LGR0,  $\sim 1.03$  for LGRs, and  $\sim 1.05$  for LGRv (not shown). Hence, the performance of our proposed method with  $K = 3$  was slightly worse than the maximum-likelihood method, completely same as the proposed method with  $K = 1$  (section 3b). Moreover, Kliche et al. (2008) computed a median value of normalized  $\lambda$  estimated with the moment method being  $\sim 1.05$  with a sample size of 1000, which is mostly the same performance as that of the proposed method with  $K = 3$ . Therefore, the proposed method can estimate the particle size distribution discriminating

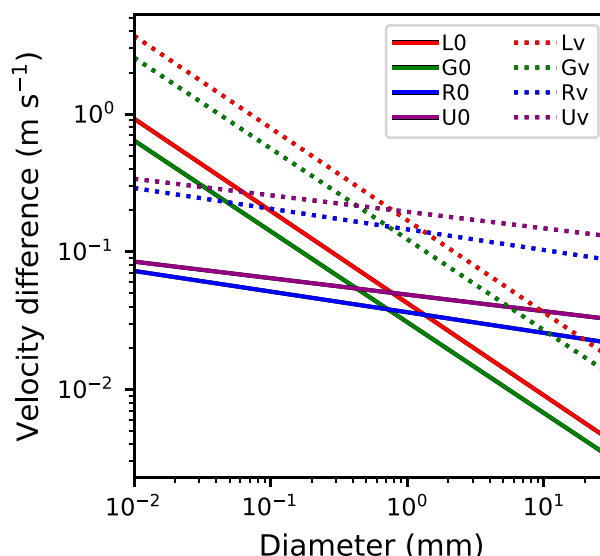


FIG. 10. The difference of mean terminal velocity between the joint PDF for a disdrometer observing particle number concentration [Eq. (5)] and the joint PDF for a disdrometer observing particles passing a finite area [Eq. (8)] expressed by  $\sigma^2 a_v D^{b_v}$  [Eqs. (2) and (10)].

PDF elements as accurately as the moment method with a single PDF.

Last, we discuss the problem that the diameter-velocity data by disdrometers such as 2DVD, PARSIVEL, and LPM originally follows the different form of joint PDF [Eq. (8)] from the joint PDF assumed in this study [Eq. (5)]: a shifted shape parameter as much as  $b_v$  [Eq. (8)] and an offset of  $\sigma^2/(a_v D^{b_v})$  in the mean terminal velocity [Eq. (9)]. The difference in terminal velocity is larger for the smaller particle diameter (Fig. 10), so that parameter  $a_v$  and  $b_v$  should be, respectively, overestimated and underestimated under an assumption of the joint PDF form by Eq. (5) comparing with the estimation assuming the joint PDF form by Eq. (8). However, the difference in the mean velocity is mostly less than  $0.1 \text{ m s}^{-1}$  with a larger diameter than 1 mm in the parameter range used in this study (Fig. 10). Therefore, the estimation of  $a_v$  and  $b_v$  might not be largely affected regardless of the joint PDF forms expressed with Eqs. (5) and (8) if the sampling data consists of large particles. If so, the shape parameter estimated by the proposed method can be simply converted to the shape parameter estimated with an assumption of the joint PDF form of Eq. (8) by an offset of  $b_v$ .

## 7. Conclusions

A new method with the EM algorithm was developed to estimate an optimal mixed joint PDF from precipitation

particles' diameter–velocity data. The method was verified by the test to retrieve the given PDF from a randomly sampling data based on a population that contains some of PDF elements: liquid particles, graupels, densely rimed aggregates, and unrimed aggregates (Fig. 1). The performance test suggested that the sample size should exceed 2000 for the computational stability (Fig. 2) and for accuracy of parameter estimation (Fig. 3). In the estimate of the number of PDF elements, minority rejection method made a better performance than maximizing AIC or BIC (Fig. 5). Our estimate made a comparable performance to the conventional least squares method (Fig. 6) and the maximum-likelihood method (section 3b) in the estimate of a single-element population. Applying this method to the observation data with 2DVD, our method was able to detect erroneous data (Fig. 7) and to distinguish liquid-phase from solid-phase precipitation particles (Fig. 8).

*Acknowledgments.* The authors thank Dr. Naoto Nakano for giving us insightful comments on a mathematical limitation on the EM algorithm. The 2DVD data used in this study were provided by Prof. Yasushi Fujiyoshi. Observation data of surface air temperature and precipitation were obtained online from the JMA. Author Katsuyama is fully supported by JSPS KAKENHI Grant 18J12196. Author Inatsu is partly supported by JSPS KAKENHI Grants (18H03819, and 19H00963); by the Social Implementation Program for Climate Change Adaptation Technology funded by the Ministry of Education, Culture, Sports, Science, and Technology; by the Environment Research and Technology Development Fund 2-1905 of the Environmental Restoration and Conservation Agency of Japan; and by Research Field of Hokkaido Weather Forecast and Technology Development (Endowed by Hokkaido Weather Technology Center Co., Ltd.). A software source code of the developed fitting algorithm will be available online (<https://humet.sci.hokudai.ac.jp/~meteo/product-e.html>).

## APPENDIX A

### Derivation of Eqs. (8) and (11)

Since the disdrometers such as 2DVD, PARSIVEL, and LPM observes precipitation particles passing a finite area, the observed diameter–velocity data should follow the joint PDF as follows (Ignaccolo and De Michele 2014):

$$P_S(V, D) = \frac{A\Delta tV}{L} N_{\text{tot}} P(V, D). \quad (\text{A1})$$

Here,  $P(V, D)$  is the joint PDF expressed with Eq. (5). The total number of precipitation particles is represented as

$$\begin{aligned} L &= A\Delta tN_{\text{tot}} \int_0^{\infty} \left[ \int_{-\infty}^{\infty} VP(V, D) dV \right] dD \\ &= A\Delta tN_{\text{tot}} a_v \lambda^{-b_v} \frac{\Gamma(\mu + b_v + 1)}{\Gamma(\mu + 1)}. \end{aligned} \quad (\text{A2})$$

Substituting Eq. (A2) into Eq. (A1), Eq. (8) is derived. Equation (11) is simply derived by rearranging Eq. (A1).

## APPENDIX B

### Initial Parameters for the EM Algorithm

The initial parameters for the EM algorithm should be reasonably determined for an accurate parameter estimation. First, we determine two velocity–diameter curves approximately passing maximum and minimum bounds of velocity on the diameter–velocity quarter plane: one is passing a point of  $(\bar{D}, \max_{1 \leq l \leq L} V_l)$  on the diameter–velocity quarter plane with  $b_v = 1$ , and another is constant at  $V = \bar{V}/4$ . These conditions are satisfied with two parameters sets of  $(a_v, b_v) = (\hat{a}_v, 1)$  and  $(a_v, b_v) = (\check{a}_v, 0)$ , where  $\hat{a}_v$  and  $\check{a}_v$  are given by Eq. (23). Then,  $K$  sets of  $(a_v, b_v)$  are equally spaced out between  $\hat{a}_v$  and  $\check{a}_v$  and between 1 and 0 as per Eq. (22). Parameters  $\sigma^2$ ,  $\mu$ , and  $\lambda$  are simply determined by the mean and variance of diameter–velocity data and are uniformly given to all the PDF elements. The mixing fraction  $\omega$  is also uniformly given.

## APPENDIX C

### Threshold for the Minority Rejection Method

The threshold value for minority rejection method is given by  $cK$ , where  $c$  is an arbitrary positive constant and  $K$  is the number of PDF elements. The threshold value by this formulation effectively reduces a possible range of  $K$  because the  $K$  estimation with the minority rejection method never exceeds an expectation value:  $cK \leq 1/K$ . Hence

$$1 \leq K \leq 1/\sqrt{c}. \quad (\text{C1})$$

This limitation is probably helpful to avoid an overfitting problem.

Figure C1 shows a hitting rate of the number of PDF elements averaging all cases of sample size ( $L = 100, 1000, 2000, 3000, 6000, 10000$ ) with the minority rejection method as a function of  $c$ . The estimation were achieved by mostly

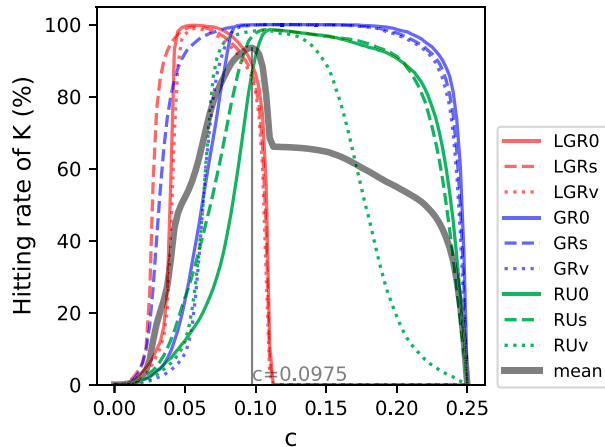


FIG. C1. Hitting rate of  $K$  by the minority rejection method as a function of  $c$ , averaged over the result of experiments with sampling data length of 100, 1000, 2000, 3000, 6000, and 10000. Line color and style is as per the legend on the right. The gray thick solid line shows the average of nine lines, with the maximum at  $c = 0.0975$ .

100% for GR0, GRs, and GRv populations within  $c$  ranging over 0.1–0.2. The number of PDF elements in RU0, RUs, and RUv populations were also successfully estimated with a relatively narrower range of  $c$ . On the other hand, for LGR0, LGRs, and LGRv populations, the PDF element number were estimated only within a narrow range of  $c$  around 0.05. The results suggested that the hitting rate becomes more sensitive to  $c$  in a case for estimating a population constructed from more number of PDF elements. Eventually, an averaged hitting rate over all population tests was approximately 90% at  $c = 0.1$  (gray thick line in Fig. C1), which limits a range of  $K$  to 1–3 [Eq. (C1)]. We hence used  $c = 0.1$  for the minority rejection method in the text.

## APPENDIX D

### Derivation of Eq. (29)

The relationship between mass and diameter of precipitation particles is empirically described by the power law as

$$M = a_m D^{b_m}, \quad (\text{D1})$$

where  $M$  (mg) is the mass of a particle and  $a_m$  and  $b_m$  are empirical coefficients (Locatelli and Hobbs 1974). Multiplying  $M$  [Eq. (D1)] and  $V$  [Eq. (2)], a mass flux  $F$  ( $\text{mg m s}^{-1}$ ) is obtained:

$$F = a_v a_m D^{b_v + b_m}. \quad (\text{D2})$$

Assuming that a PSD follows the gamma form represented by Eq. (1), precipitation intensity  $R$  ( $\text{mg m}^{-2} \text{s}^{-1}$ ) is obtained by integrating  $N_d F$  over  $D$ :

$$\begin{aligned} R &= N_0 a_v a_m \int_0^\infty D^{\mu + b_v + b_m} \exp(-\lambda D) dD \\ &= N_0 a_v a_m \frac{\Gamma(\mu + b_v + b_m + 1)}{\lambda^{\mu + b_v + b_m + 1}}. \end{aligned} \quad (\text{D3})$$

Similarly, particle number flux  $\tilde{L}$  ( $\text{m}^{-2} \text{s}^{-1}$ ) is obtained by integrating  $N_d V$  over  $D$ :

$$\tilde{L} = N_0 a_v \int_0^\infty D^{b_v} \exp(-\lambda D) dD = N_0 a_v \frac{\Gamma(\mu + b_v + 1)}{\lambda^{\mu + b_v + 1}}. \quad (\text{D4})$$

Dividing Eq. (D4) by Eq. (D3),

$$\tilde{L} = \frac{\lambda^{b_m} \Gamma(\mu + b_v + 1)}{a_v \Gamma(\mu + b_v + b_m + 1)} R. \quad (\text{D5})$$

For  $K$  elements, a linear combination of independently estimated particle number flux  $\tilde{L}_k$  following Eq. (D5) brings

$$\tilde{L} = \sum_{k=1}^K \frac{\omega_k \lambda^{b_{mk}} \Gamma(\mu_k + b_{vk} + 1)}{a_{mk} \Gamma(\mu_k + b_{vk} + b_{mk} + 1)} R. \quad (\text{D6})$$

The total particle number passing through an area  $A$  ( $\text{m}^2$ ) within a time-interval  $\Delta t$ ,  $L = A \int_0^{\Delta t} \tilde{L} d\tau$ , provides Eq. (29).

## REFERENCES

- Adirosi, E., E. Volpi, F. Lombardo, and L. Baldini, 2016: Raindrop size distribution: Fitting performance of common theoretical models. *Adv. Water Resour.*, **96**, 290–305, <https://doi.org/10.1016/j.advwatres.2016.07.010>.
- Akaike, H., 1974: A new look at the statistical model identification. *IEEE Trans. Autom. Control*, **19**, 716–723, <https://doi.org/10.1109/TAC.1974.1100705>.
- Angulo-Martínez, M., S. Beguería, B. Latorre, and M. Fernández-Raga, 2018: Comparison of precipitation measurements by OTT Parsivel<sup>2</sup> and Thies LPM optical disdrometers. *Hydrol. Earth Syst. Sci.*, **22**, 2811–2837, <https://doi.org/10.5194/hess-22-2811-2018>.
- Atlas, D., R. C. Srivastava, and R. S. Sekhon, 1973: Doppler radar characteristics of precipitation at vertical incidence. *Rev. Geophys.*, **11**, 1–35, <https://doi.org/10.1029/RG011i001p00001>.
- Battaglia, A., E. Rustemeier, A. Tokay, U. Blahak, and C. Simmer, 2010: PARSIVEL snow observations: A critical assessment. *J. Atmos. Oceanic Technol.*, **27**, 333–344, <https://doi.org/10.1175/2009JTECHA1332.1>.
- Baumgardner, D., G. Kok, W. Dawson, D. O'Connor, and R. Newton, 2002: A new ground based precipitation spectrometer: The Meteorological Particle Sensor (MPS). *11th Conf. on Cloud Physics*, Ogden, UT, Amer. Meteor. Soc., 8.6, [https://ams.confex.com/ams/11AR11CP/techprogram/session\\_12983.htm](https://ams.confex.com/ams/11AR11CP/techprogram/session_12983.htm).
- Bernauer, F., K. Hürkamp, W. Rühm, and J. Tschiersch, 2015: On the consistency of 2-D video disdrometers in measuring microphysical parameters of solid precipitation. *Atmos. Meas. Tech.*, **8**, 3251–3261, <https://doi.org/10.5194/amt-8-3251-2015>.

- , —, —, and —, 2016: Snow event classification with a 2D video disdrometer—A decision tree approach. *Atmos. Res.*, **172–173**, 186–195, <https://doi.org/10.1016/j.atmosres.2016.01.001>.
- Brandes, E. A., K. Ikeda, G. Zhang, M. Schönhuber, and R. M. Rasmussen, 2007: A statistical and physical description of hydrometeor distributions in Colorado snowstorms using a video disdrometer. *J. Appl. Meteor. Climatol.*, **46**, 634–650, <https://doi.org/10.1175/JAM2489.1>.
- Campbell, L., W. J. Steenburgh, Y. Yamada, M. Kawashima, and Y. Fujiyoshi, 2018: Influences of orography and coastal geometry on a transverse-mode sea-effect snowstorm over Hokkaido Island, Japan. *Mon. Wea. Rev.*, **146**, 2201–2220, <https://doi.org/10.1175/MWR-D-17-0286.1>.
- Cao, Q., and G. Zhang, 2009: Errors in estimating raindrop size distribution parameters employing disdrometer and simulated raindrop spectra. *J. Appl. Meteor. Climatol.*, **48**, 406–425, <https://doi.org/10.1175/2008JAMC2026.1>.
- Dempster, A. P., N. M. Laird, and D. B. Rubin, 1977: Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc.*, **39B**, 1–22, <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x>.
- Gunn, K. L. S., and J. S. Marshall, 1958: The distribution with size of aggregate snowflakes. *J. Meteor.*, **15**, 452–461, [https://doi.org/10.1175/1520-0469\(1958\)015<0452:TDWSOA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1958)015<0452:TDWSOA>2.0.CO;2).
- Handwerker, J., and W. Straub, 2011: Optimal determination of parameters for gamma-type drop size distributions based on moments. *J. Atmos. Oceanic Technol.*, **28**, 513–529, <https://doi.org/10.1175/2010JTECHA1474.1>.
- Householder, A. S., 1970: *The Numerical Treatment of a Single Nonlinear Equation*. McGraw-Hill, 216 pp.
- Huang, G. J., V. N. Bringi, R. Cifelli, D. Hudak, and W. A. Petersen, 2010: A methodology to derive radar reflectivity–liquid equivalent snow rate relations using C-band radar and a 2D video disdrometer. *J. Atmos. Oceanic Technol.*, **27**, 637–651, <https://doi.org/10.1175/2009JTECHA1284.1>.
- Ignaccolo, M., and C. De Michele, 2014: Phase space parameterization of rain: The inadequacy of gamma distribution. *J. Appl. Meteor. Climatol.*, **53**, 548–562, <https://doi.org/10.1175/JAMC-D-13-050.1>.
- Ishizaka, M., 1995: Measurement of falling velocity of rimed snowflakes (in Japanese). *Seppyo*, **57**, 229–238.
- , H. Motoyoshi, S. Nakai, T. Shina, T. Kumakura, and K. Muramoto, 2013: A new method for identifying the main type of solid hydrometers contributing to snowfall from measured size-fall speed relationship. *J. Meteor. Soc. Japan*, **91**, 747–762, <https://doi.org/10.2151/jmsj.2013-602>.
- JMA, 2019a: The past weather data (in Japanese). Accessed 24 December 2019, [https://www.data.jma.go.jp/obd/stats/etrn/view/hourly\\_s1.php?prec\\_no=14&block\\_no=47412&year=2017&month=1&day=22&view](https://www.data.jma.go.jp/obd/stats/etrn/view/hourly_s1.php?prec_no=14&block_no=47412&year=2017&month=1&day=22&view).
- , 2019b: The past weather data (in Japanese). Accessed 24 December 2019, [https://www.data.jma.go.jp/obd/stats/etrn/view/hourly\\_s1.php?prec\\_no=14&block\\_no=47412&year=2016&month=12&day=6&view](https://www.data.jma.go.jp/obd/stats/etrn/view/hourly_s1.php?prec_no=14&block_no=47412&year=2016&month=12&day=6&view).
- Johnson, R. W., D. V. Kliche, and P. L. Smith, 2014: Maximum likelihood estimation of gamma parameters for coarsely binned and truncated raindrop size data. *Quart. J. Roy. Meteor. Soc.*, **140**, 1245–1256, <https://doi.org/10.1002/qj.2209>.
- Kliche, D. V., P. L. Smith, and R. W. Johnson, 2008: L-moment estimators as applied to gamma drop size distributions. *J. Appl. Meteor. Climatol.*, **47**, 3117–3130, <https://doi.org/10.1175/2008JAMC1936.1>.
- Kruger, A., and W. F. Krajewski, 2002: Two-dimensional video disdrometer: A description. *J. Atmos. Oceanic Technol.*, **19**, 602–617, [https://doi.org/10.1175/1520-0426\(2002\)019<0602:TDVDAD>2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019<0602:TDVDAD>2.0.CO;2).
- Locatelli, J. D., and P. V. Hobbs, 1974: Fall speeds and masses of solid precipitation particles. *J. Geophys. Res.*, **79**, 2185–2197, <https://doi.org/10.1029/JC079i015p02185>.
- Löffler-Mang, M., and J. Joss, 2000: An optical disdrometer for measuring size and velocity of hydrometers. *J. Atmos. Oceanic Technol.*, **17**, 130–139, [https://doi.org/10.1175/1520-0426\(2000\)017<0130:AODFMS>2.0.CO;2](https://doi.org/10.1175/1520-0426(2000)017<0130:AODFMS>2.0.CO;2).
- Mallet, C., and L. Barthes, 2009: Estimation of gamma raindrop size distribution parameters: Statistical fluctuations and estimation errors. *J. Atmos. Oceanic Technol.*, **26**, 1572–1584, <https://doi.org/10.1175/2009JTECHA1199.1>.
- Marshall, J. S., and W. M. Palmer, 1948: The distribution of raindrops with size. *J. Meteor.*, **5**, 165–166, [https://doi.org/10.1175/1520-0469\(1948\)005<0165:TDORWS>2.0.CO;2](https://doi.org/10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2).
- Minda, H., T. Makino, N. Tsuda, and Y. Kaneko, 2016: Performance of a laser disdrometer with hydrometeor imaging capabilities and fall velocity estimates for snowfall. *IEEJ Trans. Electr. Electron. Eng.*, **11**, 624–632, <https://doi.org/10.1002/tee.22280>.
- Molthan, A. L., and B. A. Colle, 2012: Comparisons of single- and double-moment microphysics schemes in the simulation of a synoptic-scale snowfall event. *Mon. Wea. Rev.*, **140**, 2982–3002, <https://doi.org/10.1175/MWR-D-11-00292.1>.
- Muramoto, K., and T. Shiina, 1989: Measurement of snowflake size and falling velocity by image processing (in Japanese). *J. Inst. Electron. Inf. Commun. Eng.*, **J72-D-II**, 1382–1383.
- Nagumo, N., and Y. Fujiyoshi, 2015: Microphysical properties of slow-falling and fast-falling ice pellets formed by freezing associated with evaporative cooling. *Mon. Wea. Rev.*, **143**, 4376–4392, <https://doi.org/10.1175/MWR-D-15-0054.1>.
- Sasyo, Y., and T. Matsuo, 1980: On the statistical investigation of fall velocity of snowflakes. *Pap. Meteor. Geophys.*, **31**, 61–79, <https://doi.org/10.2467/mripapers.31.61>.
- Schmitt, C. G., K. Sulia, Z. J. Lebo, A. J. Heymsfield, V. Przybylo, and P. Connolly, 2019: The fall speed variability of similarly sized ice particle aggregates. *J. Appl. Meteor. Climatol.*, **58**, 1751–1761, <https://doi.org/10.1175/JAMC-D-18-0291.1>.
- Schwarz, G., 1978: Estimating the dimension of a model. *Ann. Stat.*, **6**, 461–464, <https://doi.org/10.1214/aos/1176344136>.
- Smith, P. L., and D. V. Kliche, 2005: The bias in moment estimators for parameters of drop size distribution functions: Sampling from exponential distributions. *J. Appl. Meteor.*, **44**, 1195–1205, <https://doi.org/10.1175/JAM2258.1>.
- , —, and R. W. Johnson, 2009: The bias and error in moment estimators for parameters of drop size distribution functions: Sampling from gamma distributions. *J. Appl. Meteor. Climatol.*, **48**, 2118–2126, <https://doi.org/10.1175/2009JAMC2114.1>.
- Vivekanandan, J., G. Zhang, and E. Brandes, 2004: Polarimetric radar estimators based on a constrained gamma drop size distribution model. *J. Appl. Meteor.*, **43**, 217–230, [https://doi.org/10.1175/1520-0450\(2004\)043<0217:PREBOA>2.0.CO;2](https://doi.org/10.1175/1520-0450(2004)043<0217:PREBOA>2.0.CO;2).
- Yano, J., A. J. Heymsfield, and V. T. J. Phillips, 2016: Size distributions of hydrometeors: Analysis with the maximum entropy principle. *J. Atmos. Sci.*, **73**, 95–108, <https://doi.org/10.1175/JAS-D-15-0097.1>.
- , —, and A. Bansemmer, 2018: Determination of the ice particle size distributions using observations as the integrated constraints. *J. Atmos. Sci.*, **75**, 787–804, <https://doi.org/10.1175/JAS-D-17-0145.1>.
- Yuter, S. E., D. E. Kingsmill, L. B. Nance, and M. Löffler-Mang, 2006: Observations of precipitation size and fall speed characteristics within coexisting rain and wet snow. *J. Appl. Meteor. Climatol.*, **45**, 1450–1464, <https://doi.org/10.1175/JAM2406.1>.