Improving the Representation of Raindrop Size Distributions Using the In-situ Microphysics Observations Collected in Hurricanes

Hua Leighton[1], Xuejin Zhang[2], Robert A. Black[3], and Frank Marks[2]

[1]University of Miami
[2]NOAA/AOML
[3]NOAA/AOML/HRD

December 22, 2022

Abstract

Raindrop Size Distributions (RSDs) samples from 17 flight missions though 6 hurricanes collected by Precipitation Imaging Probe (PIP) during National Oceanic and Atmospheric Administration's hurricane field program in 2020 are used to study gamma fits of the RSDs in hurricanes. The method of moment (MM) is adopted for solving for the three parameters in gamma distribution. The results show that the usage of lower (higher) moments produces large biases for integral rain variables (IRV) of higher (lower) moments. These biases can be alleviated by extracting the best fits from five groups that use increasing higher orders of moments for MM. An intercept (N_0)— slope (λ) relation identified from the fitted gamma distributions captures 92% of the variance of the data, where the majority of remaining 8% can be further captured by including the impact of liquid water content (LWC), as shown in the results from a random forest regression model.

Hosted file

951453_0_art_file_10556148_rn9bvm.docx available at https://authorea.com/users/568909/articles/614551-improving-the-representation-of-raindrop-size-distributions-using-the-in-situ-microphysics-observations-collected-in-hurricanes

1

Improving the Representation of Raindrop Size Distributions Using the In-situ

Microphysics Observations Collected in Hurricanes

Hua Leighton[1,2], Xuejin Zhang[2], Robert Black[2], Frank D. Marks, Jr.[2]

[1]Rosenstiel School of Marine and Atmospheric Science, University of Miami

[2]NOAA/AOML/Hurricane Research Division, Miami, Florida

Key Points:

- Method of moments produces bias when fitting raindrop size distribution (RSDs) but the bias can be alleviated by composite moment fitting.

- The identified $N_0 - \lambda$ relation captures 92% of the variance in the fitted RSDs that have correlation coefficients larger than 0.9.

- A random forests regression model taking both N_0 and Liquid Water Content as inputs captures most of remaining 8% of variance in the data.

16   **Abstract**

17   Raindrop Size Distributions (RSDs) samples from 17 flight missions though 6 hurricanes

18   collected by Precipitation Imaging Probe (PIP) during National Oceanic and Atmospheric

19   Administration's hurricane field program in 2020 are used to study gamma fits of the RSDs in

20   hurricanes. The method of moment (MM) is adopted for solving for the three parameters in

21   gamma distribution. The results show that the usage of lower (higher) moments produces large

22   biases for integral rain variables (IRV) of higher (lower) moments. These biases can be

23   alleviated by extracting the best fits from five groups that use increasing higher orders of

24   moments for MM. An intercept (N_0) — slope (λ) relation identified from the fitted gamma

25   distributions captures 92% of the variance of the data, where the majority of remaining 8% can

26   be further captured by including the impact of liquid water content (LWC), as shown in the

27   results from a random forest regression model.

28   **Plain Language Summary**

29   How well an assumed statistical distribution can represent the number of raindrops in each size

30   bin is crucial to both accurate rainfall estimation from observed radar echo and successful

31   forecasts of numerical weather models. Gamma distribution, one of statistical distributions, is

32   often used and the three parameters (i.e. intercept, slope and shape) of gamma distribution are

33   obtained by solving three equations. Different set of three equations can lead to different

34   solutions, each of which has its advantage and disadvantage. In this study, we explore five

35   different sets of three equations, extract the solutions that have low bias and high correlation

36   coefficient from each set, and develop composite solutions. We investigate the relationships

37   between each pair of the three parameters for the composite solutions and find intercept and

38   slope are closely related. A linear fit that represents intercept-slope relationship very well already

39   is further improved by using a machine learning model that takes into account both intercept and

40   the mass of raindrops.

41

42   **1 Introduction**

43   With the rapid advancement of computational technology, numerical models have become the

44   most important tool in forecasting hurricane intensity and precipitation. In operational numerical

45   models, bulk microphysics parameterization schemes are used due to their computational

46   efficiency. The bulk schemes assume the size distributions of each hydrometeor category to be

47   certain statistical distribution. The formulations of all the microphysical processes can then be

48   derived from these assumed statistical distributions and other assumptions made in the scheme.

49   The microphysical processes play a significant role in the distribution of diabatic heating, which

50   is one of the primary driving forces of a tropical cyclone's intensity change. The realistic

51   representation of microphysical processes in numerical models is crucial to simulating the

52   intensity and structure evolution of hurricanes accurately. Early studies (e.g., Marshall and

53   Palmer 1948; Mueller and Sims 1966; Sulakvelidze 1969) have proposed many different

54   statistical distributions representing RSDs. Among them the gamma distribution has been widely

55   used due to their generalized representation for RSD. The gamma distribution,

56   $$N(D) = N_0 D^\mu e^{-\lambda D} \quad \text{(eq. 1)}$$

57   as shown above, with three parameters, intercept N_0, shape parameter μ, and slope λ, is able to

58   adequately describe the small spatiotemporal-scale variations of RSDs in most situations

59   (Ulbrich and Atlas 1998). It reduces to the exponential distribution when the shape parameter μ

60   is zero. The gamma distribution also makes it particularly easy to calculate the moments and

61   formulate microphysical processes in the bulk schemes. The original interest of RSD studies is to

62   estimate IRVs, such as rainfall (e.g., Seliga and Bringi 1976; Gorgucci et al. 1994; Ulbrich and

63   Atlas 1998). As stated in Kozu and Nakamura (1991), assuming RSDs to be a two- or three-

64   parameter statistical distribution, measuring two or three IRVs can determine the RSD

65   parameters, thereby enabling an accurate estimation of other IRVs. For this purpose, the method

66   of moments (MM) has been widely used. However, studies (Haddad et al. 1996, 1997; Smith and

67   Kliche 2005; Smith et al. 2009) pointed out MM produced biases. These biases might not have

68   significant impact on the application of estimating certain IRVs but can drastically change the

69   outcome of microphysics processes that are formulated based on the fitted RSDs. Therefore, for

70   modeling purposes, it is crucial to minimize the biases while still maintain the accuracy of the

71   calculated IRVs.

72   The data used this study is briefly introduced in section 2. The remainder of the paper is

73   organized as follows. In section 3, the results of the gamma fitting to the RSDs are presented.

74   The identified $N_0 - \lambda$ relation is presented in section 4. An improved $N_0 - \lambda$ relation using

75   random forests (RF) regression model is demonstrated in section 5 and followed by a discussion

76   and conclusion section.

77   **2. Data**

78   The RSD observations are from National Oceanic and Atmospheric Administration's hurricane

79   field program in 2020. They were collected by the Droplet Measurement Technologies (DMT)

80   Precipitation Imaging Probe in 6 hurricanes from 17 flights, i.e., 4 flights from Hanna, 3 flights

81   from Isaias, 2 flights from Laura, 3 flights from Sally, 1 flight from Zeta, and 4 flights from

82   Delta. Observations taken within the 500-km radius of the storm center are included. The number

83   of total RSD observations used is 18076 in this study. The detailed description of the data set is

84   provided in Leighton et al. (2022)

85   **3 Methodology**

86   The moment of a raindrop size distribution is defined as:

87
$$M_m = \int D^m N(D) dD \qquad \text{(Eq. 2)}$$

88   Where m is the number of moments, $N(D)$ is the raindrop size distribution as the function of

89   diameter D. Inserting Eq. 1 into Eq.2 above, we arrive at

90
$$M_m = N_0 \frac{\Gamma(m+\mu+1)}{\lambda^{m+\mu+1}} \qquad \text{(Eq. 3)}$$

91   Given the special property of gamma function,

92
$$\Gamma(\alpha + 1) = \alpha \, \Gamma(\alpha) \qquad \text{(Eq. 4)}$$

93   any group of three consecutive moments gives a set of unique solutions of the three parameters

94   for the gamma distribution. After manipulating Eq.3 for three consecutive moments (e.g. m, m+1

95   and m+2), we obtain the solutions as following,

96
$$\mu = \frac{Cm + C - m - 2}{1 - C} \qquad \text{(Eq. 5)}$$

97
$$\lambda = B(m + \mu + 1) \qquad \text{(Eq. 6)}$$

98
$$N_0 = \frac{M_m \lambda^{m+\mu+1}}{\Gamma(m+\mu+1)} \qquad \text{(Eq. 7)}$$

99 Where C is $\frac{M_m M_{m+2}}{M_{m+1}^2}$ and B is $\frac{M_m}{M_{m+1}}$. $M_m, M_{m+1}, and\ M_{m+2}$ are three consecutive moments

100 calculated from Eq. 2 where $N(D)$ is the observed RSDs.

101 In this study, we explore five different combinations of three consecutive moments, i.e. moments

102 0, 1, 2 (m012), moments 1, 2, 3 (m123), moments 2, 3, 4 (m234), moments 3, 4, 5 (m345), and

103 moments 4, 5, 6 (m456). It is worth noting that the moments above are calculated directly from

104 the observed RSDs and the calculated moments might not correspond to the IRVs of the same

105 moments. For example, the 3rd moment calculated above is not the same as LWC since the

106 density of water is not taken into account. The performance of gamma fits is evaluated from two

107 aspects: 1) comparing the IRVs calculated from the fitted gamma distributions and that from the

108 observed RSDs, and 2) comparing RSD shapes by calculating the correlation coefficient between

109 the fitted RSD and the observed RSD. The five IRVs used for evaluating the performance of the

110 fitted RSDs are total number of concentrations, mass-weighted-diameter, LWC, radar reflectivity

111 and rainfall rate. The calculations of these IRVs are shown in the following from equations 5-9:

112
$$N_t = \int_0^\infty N(D)dD \qquad \text{(Eq. 8)}$$

113
$$D_m = \frac{\int_0^\infty D^4 N(D)dD}{\int_0^\infty D^3 N(D)dD} \qquad \text{(Eq. 9)}$$

114
$$LWC = \frac{\pi}{6}\rho_w \int_0^\infty D^3 N(D)\ dD \qquad \text{(Eq. 10)}$$

115
$$Ref = 10\ log_{10} \int_0^\infty D^6 N(D)dD \qquad \text{(Eq. 11)}$$

116
$$RR = \frac{\pi}{6}\rho_w \int_0^\infty D^3 N(D)\ V_t(D)dD \qquad \text{(Eq. 12)}$$

117   ρ_w in Eq. (7) and (9) is the density of water, $1000\ kg\ m^{-3}$. V_t in Eq. (12) is the terminal velocity

118   of raindrops and follows Best (1950).

4 Results

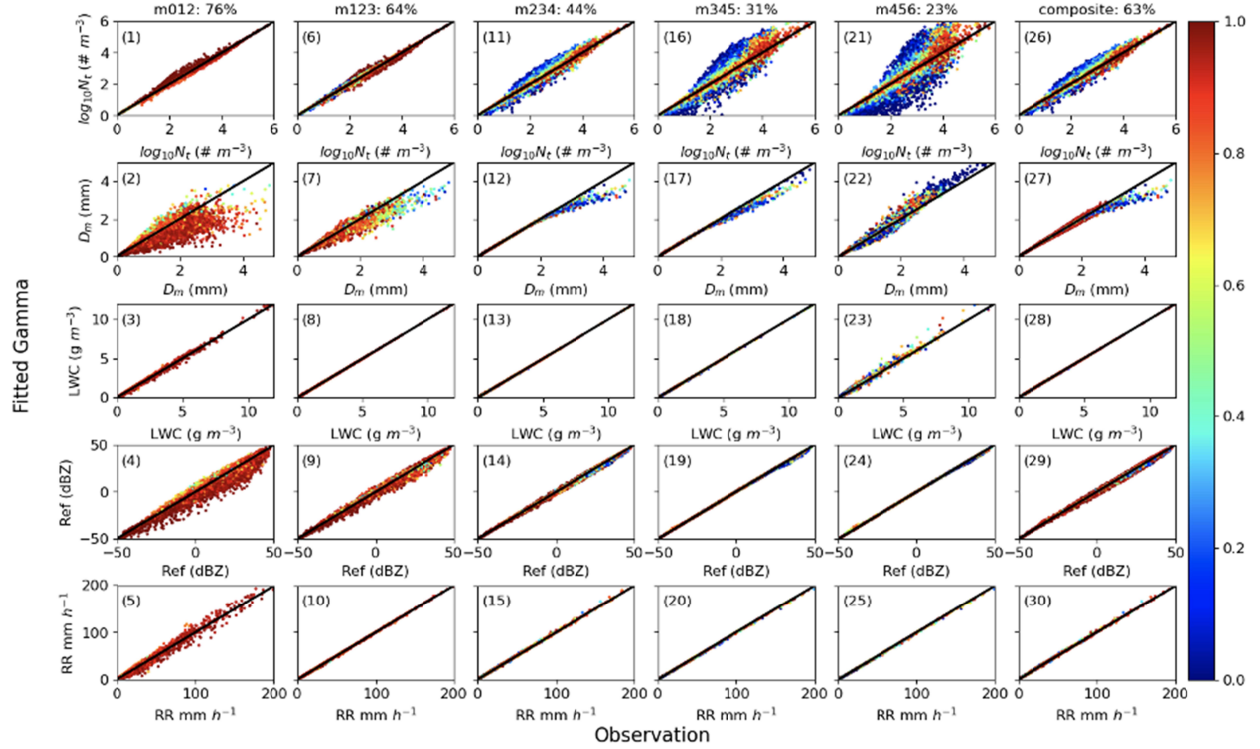### 4.1 Gamma fitting of RSDs

121
122   Figure 1: Scatter plots of different IRVs (i.e. total number concentration N_t, mass-weighted-

123   diameter D_m, liquid water content LWC, radar reflectivity Ref, and rainfall rate RR) calculated

124   from the observed RSDs (abscissa) and the fitted gamma distributions (ordinate). The first five

125   columns are from moments 0, 1 and 2 (m012), moments 1, 2, and 3 (m123), moments 2, 3 and 4

126   (m234), moments 3, 4 and 5 (m345) and moments 4, 5 and 6 (m456), respectively. Last column,

127   termed as composite, are obtained by merging the fitted gamma distribution from all five groups

128   of three consecutive moments such that the best fits of each group are retained (see text for

129   details). The corresponding moments and the percentage of fits that have correlation

130   coefficient >0.9 are shown at the top of each column. The color indicates the correlation

131   coefficient between the fitted gamma distribution and the observed distribution. The black line in

132   each plot shows where the IRVs calculated from the fitted gamma distributions equals that

133   calculated from the observed distributions.

134   As shown in Fig. 1, the IRVs calculated from the gamma fitting that uses three lowest moments

135   (m012) have the highest correlation coefficients, and 76% of fits having correlation

136   coefficients >0.9. This ratio decreases with the increasing order of moments and is only 23% for

137   m456. This is consistent with Smith and Kliche (2005) and Smith et al. (2009), who showed that

138   errors of the estimates of the RSD parameters using MM are usually larger when higher-order

139   moments are employed and suggested using the lowest-order sample moments. However, when

140   evaluated by the calculated IRVs, the MM using the lowest-order moments has the worst

141   performance for most of the calculated IRVs. Figure 1 (panel 1-5) shows that m012 slightly

142   overestimates the total number concentration but severely underestimates mass-weighted-

143   diameter, especially for the large drops. Consequently, radar reflectivity is drastically

144   underestimated in m012. The rainfall rate is also underestimated for rainfall rates <100 mm hr^{-1}.

145   The LWC from m012 is in good agreement with the observation. As the order of moments

146   increases to 3, 4 and 5 in Fig 1 (panel 16-20)), the calculation of higher moments reaches the

147   optimum. The LWC, reflectivity and rainfall rate all agree well with the observation. Yet the

148   number concentration degrades significantly, even compared to m234. As the order of moments

149   increases to 4, 5 and 6, mass-weighted-diameter is mostly overestimated and so is the LWC.

150   However, this overestimation is offset by the underestimation of number concentration and

151   consequently both reflectivity and rainfall rate showed good agreement with the observation.

152   This highlights the deficiency of evaluating the fits based on only one IRV. For example, Tokay

153   and Short (1996) showed that calculated rainfall rates from fitted distributions are in excellent

154   agreement with rainfall rates obtained from observed RSDs in their Fig. 1. However, the

155   excellent agreement for rainfall rates alone does not guarantee that fitted distributions represent

156   the observed RSDs well. As seen in Fig. 1, lower order moments produce good agreement with

157   the observations for IRVs such as total number concentration, and higher order moments produce

158   good agreement with the observations for IRVs such as reflectivity and rainfall rate. Middle

159   moments, such as m234, show a good balance that generates overall good agreement for the

160   calculated IRVs. Therefore, our approach is to composite the moment fits from all five groups so

161   that best fits of each group can be utilized. The composite is compiled according to the following

162   approach. First, the fits from M456 are selected if their correlation coefficient is >0.9 and the

163   error of total number concentration is <10 m^{-3}. The same selection criteria are applied to the

164   remaining samples (total samples minus selected samples from M456), but the fits are selected

165 from M345. Next, for the updated remaining samples (total samples minus the selected samples

166 from M456 and minus the selected samples from M345), the fits from M234 are selected if their

167 correlation coefficient is >0.9. For the new remaining samples (total samples - selected samples

168 from M456 - selected samples from M345 - selected samples from M234), the fits from M123

169 are selected if their correlation coefficient is >0.9 and the error of D_m is <10%. The same

170 screening process is performed for the remaining samples, but the fits are selected from M012.

171 The rest of the fits come from M234. The overall pattern of this composite MM fitting will

172 resemble M234 but the correlation coefficient is expected to improve. As shown in the last

173 column of Fig. 1 (panel 26-30), the calculated IRVs agree well with the observation in general,

174 as in M234. Yet the ratio of fits that have correlation coefficients >0.9 increases from 44% from

175 M234 to 63% in the composite MM. The distributions of the three parameters for the fits with

176 correlation coefficient >0.9 from the composite MM are shown in Fig. 2.

177 **4.2 $N_0 - \lambda$ relationship from fitted gamma distributions**

178 Many studies (e.g. Ulbrich 1983; Zhang et al. 2001, 2003; Brandes et al. 2003, 2004;

179 Vivekanandan et al. 2004; Ulbrich and Atlas 2007) have explored the relationships among the

180 three parameters in the fitted gamma distributions of RSDs. Ulbrich (1983) showed the

181 relationship between N_0 and μ that is deduced from empirical relations between IRVs, such as Z-

182 R relationship, from early studies. Other studies (Zhang et al. 2001, 2003; Brandes et al. 2003,

183 2004; Vivekanandan et al. 2004; Ulbrich and Atlas 2007, Chang et al. 2009) deduced $\mu - \lambda$

184 relationships based on fitted RSDs from different data sources. They show that this relationship

185 provides useful information to describe RSDs and improves the accuracy of the retrieved RSDs

186 from polarimetric radar measurements. The relationship between any pair of parameters can also

187 be used in microphysics parameterization schemes. For a one-moment scheme that uses the

188 gamma distribution for RSDs, when one parameter is prescribed, the second parameter can be

189 calculated from the relationship between this pair, and the third parameter can be diagnosed from

190 the prognostic variable LWC and the two known parameters. For a two-moment scheme, two

191 prognostic variables (e.g. LWC and the number concentration) and the relationship can fully
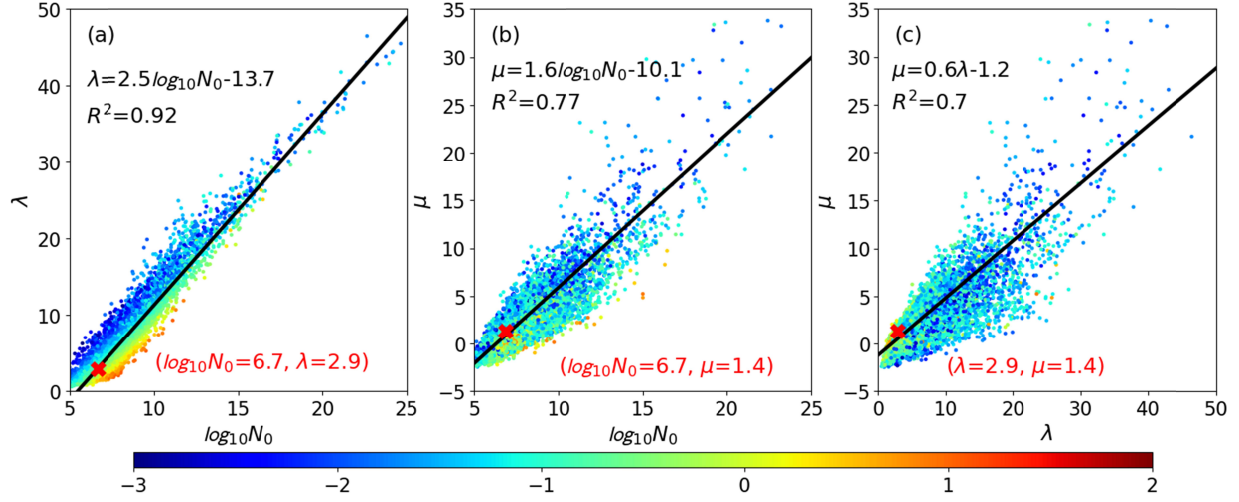
192 determine the RSDs.

193

Figure 2: Scatter plots of a) Slope λ vs. $log_{10}N_0$ b) Shape μ vs. $log_{10}N_0$, and c) Shape μ vs. Slope λ. The data points are colored by log_{10} (LWC). The red cross indicates the mode for the pair of parameters and the corresponding mode values are denoted in the red text at the bottom of each figure. The black line in each plot indicates the best fitted line and the corresponding equation is shown above the line. R^2 for the fit is shown below the equation.

Despite the $N_0 - \mu$ relation and $\mu - \lambda$ relation identified in the literature, Fig.2 shows that both the $N_0 - \mu$ (Fig. 2b) distribution and the $\mu - \lambda$ distribution (Fig. 2c) have large scatter, which makes empirical relationships for these two pairs less representative. In contrast, the scatter for the $N_0 - \lambda$ distribution (Fig. 2a) is much smaller. The linear fit (the black line in Fig. 2a) captures 92% of variance of the samples. Most of the remaining variance can be further captured by including the impact of LWC. As seen in Fig. 2a, for a given λ, N_0 increases with LWC. LWC is a prognostic variable in all bulk microphysics schemes, providing a dynamic relationship that can be used to improve the performance of the bulk microphysics schemes.

5 Improved $N_0 - \lambda$ relationship using a random forest regression model

Over the last few years, applications of machine learning in weather and climate fields have grown exponentially (e.g. Gagne II et al. 2017, McGovern et al. 2019). Random Forests (Breiman 2001) is a simple but powerful machine learning application (Gagne II et al.. 2014; Herman and Schumacher 2018). RFs learn an ensemble of decision trees, each of which is

212   trained on a separate bootstrap resampled dataset and using a different subset of the attributes. In

213   this study, we use two attributes, $log_{10}N_0$ and LWC, for the RFs input. Output is slope λ. One

214   common problem in machine learning application is overfitting. In order to objectively evaluate

215   the performance of a machine learning model, the data are always split into training data and

216   testing data. We use 70% of data for training and 30% for testing. The performance of both

217   linear regression and RFs regression is evaluated. As shown in Fig. 3a, R^2 of linear regression

218   model for testing data is 0.91, similar to that in Figure 2a. R^2 of RF regression model for testing

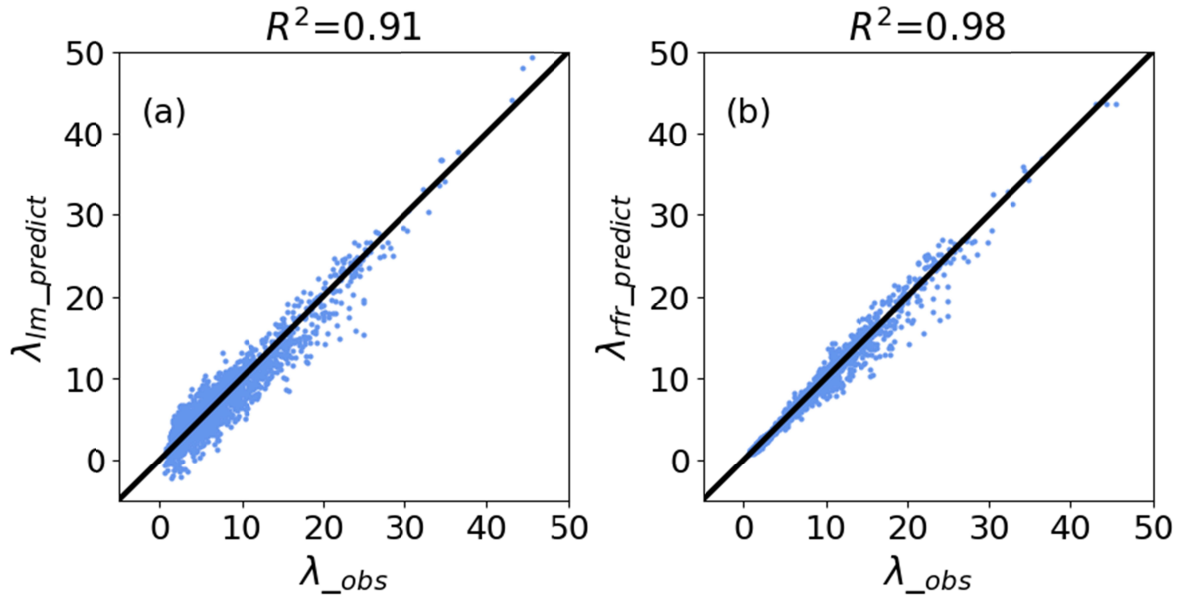219   data is 0.98, 7% increased on R^2 obtained from the linear regression model.

220
221   Figure 3: Scatter plots of a) Slope λ from composite MM fitting vs. slope λ predicted from linear

222   regression model (input: $log_{10}N_0$ from composite MM fitting), and b) Slope λ from composite

223   MM fitting vs. slope λ predicted from random forests regression model (input: $log_{10}N_0$ from

224   composite MM fitting and LWC)

225   **6 Discussion and Conclusions**

226   Raindrop Size Distributions collected by PIP from 17 flights through 6 hurricanes during

227   hurricane field program in 2020 are used to study gamma fits in hurricanes. The results from

228   gamma fitting using MM showed that using the lowest orders of moments produces the best fits

229   when evaluated by the correlation coefficient between the fitted and the observed RSDs. Yet, the

230   IRVs, especially radar reflectivity and rainfall rate, are significantly underestimated due to the

231   underestimated mass-weighted-diameter. In contrast, radar reflectivity and rainfall rate

232   calculated from high order MM fits are in excellent agreement with the observations. This

233   excellent agreement is the result of overestimated mass-weighted-diameter and underestimated

234   total number concentration, especially for large drops. The correlation coefficient is much lower

235   for high order MM fits. The central moments, M234, shows overall good performance, yet only

236   44% of fits represent the observed RSD well, evaluated by the correlation coefficient between

237   the fitted and the observed RSDs. By compiling composite MM fits to extract best fits in each

238   group, the ratio of fits with correlation coefficient larger than 0.9 increased from 44% to 63%

239   without compromising the calculated IRVs.

240   The distribution of the intercept N_0 and slope λ showed a strong correlation. A linear empirical

241   relationship that is obtained by fitting the entire dataset captures 92% of the variance of the data.

242   The remaining 8% of variance is shown to be closely related to LWC. A RF regression model is

243   able to capture 98% of the variance of the data if inputs include both N_0 and LWC. The

244   distributions of $\mu - N_0$ and $\mu - \lambda$ also show correlation in each pair (Fig. 2b and Fig. 2c) but

245   the scatter is significantly larger than that in $N_0 - \lambda$, making a fitted empirical relationship less

246   representative.

247   The identified $N_0 - \lambda$ relationship obtained from the RF regression model can not only improve

248   the accuracy of the retrieved RSDs from polarimetric radar measurements by providing useful

249   information to describe RSDs but also reduce the uncertainties and increase the accuracy of bulk

250   microphysics parameterization schemes in numerical models. For a one-moment scheme that

251   uses gamma distribution for RSDs, if N_0 is provided, then λ can be calculated from this $N_0 - \lambda$

252   relationship with high confidence and the shape parameter μ can be diagnosed from N_0, λ, and

253   the prognostic variable LWC. For a two-moment scheme, the $N_0 - \lambda$ relationship along with two

254   prognostic variables, LWC and total number concentration, can fully determine the RSD. The

255   accuracy of the microphysics processes in the bulk scheme can therefore be better formulated

256   and so potentially improves the overall performance of the microphysical parameterization

257   schemes.

258   **Acknowledgments**

265   **Data Availability Statement**

266   The authors express thanks to NOAA/HRD Data Support for providing the microphysics

267   observation data (https://www.aoml.noaa.gov/ftp/hrd/data/cloudphysics/2020/)

268

269   **References**

270   Best, A.C. (1950), Empirical formulae for the terminal velocity of water drops falling through

271   the atmosphere. *Q. J. R. Meteor. Soc.* 76, 302–311.

272   Brandes, E. A., Zhang, G. & Vivekanandan, J. (2003), An evaluation of a drop distribution–

273   based polarimetric radar rainfall estimator. *J. Appl. Meteor.*, 42, 652–660.

274   Brandes, E. A., Zhang, G. & Vivekanandan, J. (2004), Drop size distribution retrieval with

275   polarimetric radar: Model and application. *J. Appl. Meteor.*, 43, 461–475.

276   Breiman, L. (2001), Random forests, *Mach. Learn.* 45(1), 5–32.

277   Chang, W.-Y., Wang, T.-C. C., & Lin, P.-L. (2009), Characteristics of the raindrop size

278   distribution and drop shape relation in typhoon systems in the western Pacific from the 2D video

279   disdrometer and NCU C-band polarimetric radar. *J. Atmos. Oceanic Technol.*, 2610, 1973–1993.

280   Gagne II, D. J., McGovern, A., Haupt, S. E., Sobash, R., Williams, J. K. & Xue, M.

281   (2017), Storm-Based Probabilistic Hail Forecasting with Machine Learning Applied to

282   Convection-Allowing Ensembles. *Wea. and Forecasting*, 32, 1819-1840.

283   Gagne II, D. J., A. McGovern, & M. Xue, 2014: Machine learning enhancement of storm-scale

284   ensemble probabilistic quantitative precipitation forecasts. *Wea. Forecasting*, **29**, 1024–1043.

285 Gorgucci, E., and Scarchilli, G. & Chandrasekar, V. (1994), A robust estimator of rainfall rate

286 using differential reflectivity. J. Atmos. Oceanic Technol., 11, 586–592.

287 Haddad, Z. S., Durden, S. L., & Im, E., (1996), Parameterizing the raindrop size distribution. *J.*

288 *Appl. Meteor.*, 35, 3–13.

289 Haddad, Z. S., Short, D.A., Durden, S. L., Im, E., Hensley, S., Grable, M. B., & Black, R. A.

290 (1997), A new parametrization of the rain drop size distribution. *IEEE Transactions on*

291 *Geoscience and Remote Sensing*, 35, 532–539.

292 Herman, G. R., & Schumacher, R. S. (2018), "Dendrology" in numerical weather prediction:

293 What random forests and logistic regression tell us about forecasting extreme precipitation. *Mon.*

294 *Wea. Rev.*, **146**, 1785–1812

295 Kozu, T., & Nakamura, K. (1991), Rainfall parameter estimation from dual-radar measurements

296 combining reflectivity profile and path-integrated attenuation. *J. Atmos. Oceanic Technol.*, 8,

297 259–270.

298 Leighton, H., Black, R., Zhang, X., & Marks, F. D. (2022). The relationship between reflectivity

299 and rainfall rate from rain size distributions observed in hurricanes. *Geophysical Research*

300 *Letters*, 49, e2022GL099332. https://doi.org/10.1029/2022GL099332

301 McGovern, A., Gagne II, D.J., Lagerquist, R., Elmore, K., & Jergensen, G.E. (2019) Making the

302 black box more transparent: Understanding the physical implications of machine learning.

303 *Bulletin of the American Meteorological Society*, 100(11), 2175-2199.

304 Marshall, J. S., & Palmer, W. M. (1948), The distribution of raindrops with size. *J. Meteor.*, 5,

305 165-166.

306 Mueller, E. A., & Sims, A. L. (1966), Radar cross sections from drop size spectra. Tech. Rep.

307 ECOM-00032-F, Contract DA- 28-043 AMC-00032€ Illinois State Water Survey, Urbana,110

308 pp. [AD-645218].

309   Seliga, T. A., & Bringi, V. N. (1976), Potential use of the radar differential reflectivity

310   measurements at orthogonal polarizations for measuring precipitation. *J. Appl. Meteor.*, 15, 69–

311   76.

312   Smith, P. L., & Kliche, D. V. (2005), The Bias in moment estimators for parameters of drop size

313   distribution functions: Sampling from exponential distributions. *J. Appl. Meteor.*, 44, 1195–

314   1205.

315   Smith, P. L., Kliche, D. V. & Johnson, R. W. (2009), The bias and error in moment estimators

316   for parameters of drop size distribution functions: Sampling from gamma distributions. *J. Appl.*

317   *Meteor Climatol.*, **48**, 2118–2126.

318   Sulakvelidze, G. K. (1969), Rainstorms and Hail (translated from Russian), Israel Program for

319   Scientific Translations, Jerusalem [available from NTIS].

320   Tokay, A., & D. A. Short. (1996), Evidence from tropical raindrop spectra of the origin of rain

321   from stratiform versus convective clouds. *J. Appl. Meteor.,* 35, 355–371.

322   Ulbrich, C. W., (1983), Natural variations in the analytical form of the raindrop size distribution.

323   *J. Climate Appl. Meteor.*, 22, 1764–1775.

324   Ulbrich, C. W., & Atlas, D. (1998) Rainfall microphysics and radar properties: Analysis methods

325   for drop size spectra. *J. Appl. Meteor.*, 37, 912–923.

326   Ulbrich, C. W., & Atlas, D. (2007), Microphysics of raindrop size spectra: Tropical continental

327   and maritime storms. *J. Climate Appl. Meteor.*, 46, 1777–1791.

328   Vivekanandan, J., Zhang, G. & Brandes, E. (2004), Polarimetric radar estimators based on a

329   constrained gamma drop size distribution model. *J. Appl. Meteor.*, 43, 217–230.

330   Zhang, G., Vivekanandan, J., & E. Brandes (2001), A method for estimating rain rate and drop

331   size distribution from polarimetric radar measurements. *IEEE Trans. Geosci. Remote Sens.*, 39,

332   830–841.

333 Zhang, G., Vivekanandan, J., & E. Brandes, Meneghini, R., & T. Kozu, 2003: The shape – slope

334 relation in observed gamma raindrop size distribution: Statistical error or useful information? *J.*

335 *Atmos. Oceanic Technol.*, 20, 1106–1119.

336
337