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

June 11, 2023

Abstract

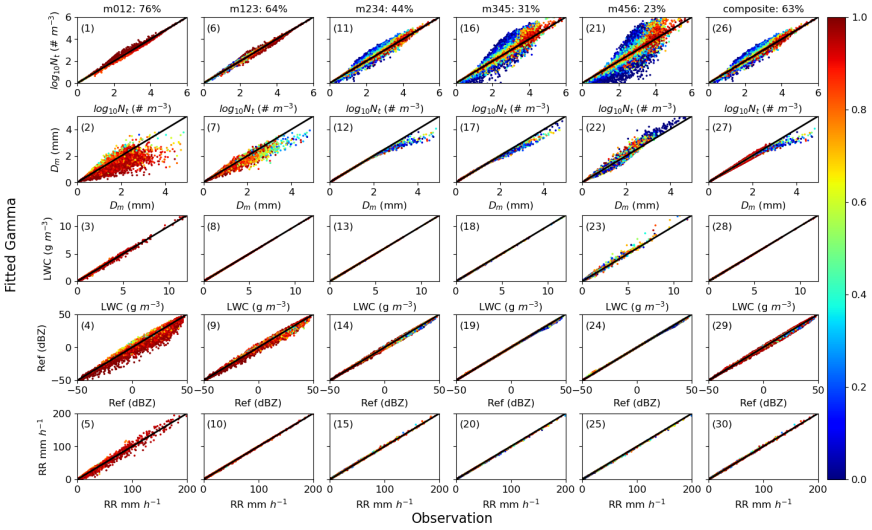
Raindrop Size Distributions (RSDs) samples from 15 flight missions though 6 hurricanes collected by Precipitation Imaging Probe (PIP) during National Oceanic and Atmospheric Administration's hurricane field program in 2020 are used to study gamma fits of the RSDs in hurricanes. The method of moment (MM) 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 (N0)— 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.

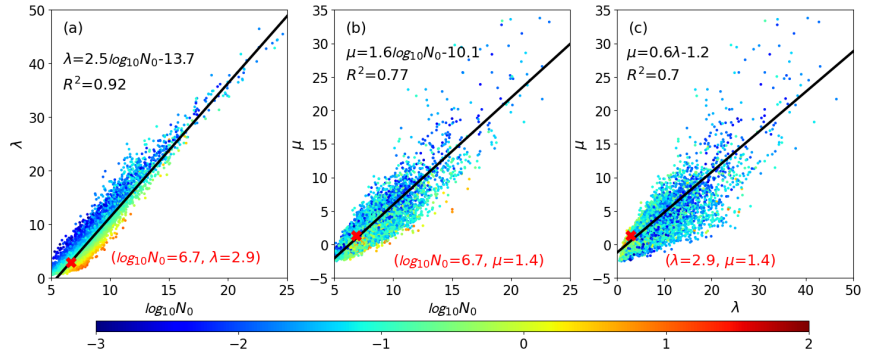
Hosted file

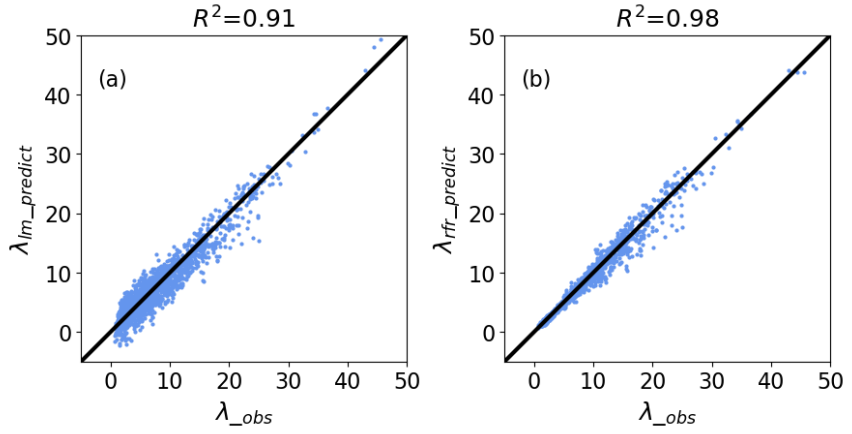
963870_0_art_file_11062326_rvc9z0.docx available at https://authorea.com/users/568909/articles/648054-improving-the-representation-of-raindrop-size-distributions-using-the-in-situ-microphysics-observations-collected-in-hurricanes

Hosted file

963870_0_supp_11062330_rvjbvs.docx available at https://authorea.com/users/568909/articles/648054-improving-the-representation-of-raindrop-size-distributions-using-the-in-situ-microphysics-observations-collected-in-hurricanes







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.

Abstract

Raindrop Size Distributions (RSDs) samples from 15 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.

Plain Language Summary

How well an assumed statistical distribution can represent the number of raindrops in each size bin is crucial to both accurate rainfall estimation from observed radar echo and successful forecasts of numerical weather models. Gamma distribution, one of statistical distributions, is often used and the three parameters (i.e. intercept, slope and shape) of gamma distribution are obtained by solving three equations. Different set of three equations can lead to different solutions, each of which has its advantage and disadvantage. In this study, we explore five different sets of three equations, extract the solutions that have low bias and high correlation coefficient from each set, and develop composite solutions. We investigate the relationships between each pair of the three parameters for the composite solutions and find intercept and slope are closely related. A linear fit that represents intercept-slope relationship very well already is further improved by using a machine learning model that takes into account both intercept and the mass of raindrops.

1 Introduction

With the rapid advancement of computational technology, numerical models have become the most important tool in forecasting hurricane intensity and precipitation. In operational numerical models, bulk microphysics parameterization schemes are used due to their computational efficiency. The bulk schemes assume the size distributions of each hydrometeor category to be certain statistical distribution. The formulations of all the microphysical processes can then be derived from these assumed statistical distributions and other assumptions made in the scheme. The microphysical processes play a significant role in the distribution of diabatic heating, which is one of the primary driving forces of a tropical cyclone's intensity change. The realistic representation of microphysical processes in numerical models is crucial to simulating the intensity and structure evolution of hurricanes accurately. Early studies (e.g., Marshall and Palmer 1948; Mueller and Sims 1966; Sulakvelidze 1969) have proposed many different statistical distributions representing RSDs. Among them the gamma distribution has been widely used due to their generalized representation for RSD. The gamma distribution,

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

as shown above, with three parameters, intercept N_0, shape parameter μ, and slope λ, is able to adequately describe the small spatiotemporal-scale variations of RSDs in most situations (Ulbrich and Atlas 1998). It reduces to the exponential distribution when the shape parameter μ is zero. The gamma distribution also makes it particularly easy to calculate the moments and formulate microphysical processes in the bulk schemes. The original interest of RSD studies is to estimate IRVs, such as rainfall (e.g., Seliga and Bringi 1976; Gorgucci et al. 1994; Ulbrich and Atlas 1998). As stated in Kozu and Nakamura (1991), assuming RSDs to be a two- or three-parameter statistical distribution, measuring two or three IRVs can determine the RSD parameters, thereby enabling an accurate estimation of other IRVs. For this purpose, the method of moments (MM) has been widely used. However, studies (Haddad et al. 1996, 1997; Smith and Kliche 2005; Smith et al. 2009) pointed out MM produced biases. These biases might not have significant impact on the application of estimating certain IRVs but can drastically change the outcome of microphysics processes that are formulated based on the fitted RSDs. Therefore, for modeling purposes, it is crucial to minimize the biases while still maintain the accuracy of the calculated IRVs.

The data used this study is briefly introduced in section 2. The remainder of the paper is organized as follows. In section 3, the results of the gamma fitting to the RSDs are presented. The identified $N_0 - \lambda$ relation is presented in section 4. An improved $N_0 - \lambda$ relation using random forests (RF) regression model is demonstrated in section 5 and followed by a discussion and conclusion section.

2. Data

The RSD observations are from National Oceanic and Atmospheric Administration's hurricane field program in 2020. They were collected by the Droplet Measurement Technologies (DMT) Precipitation Imaging Probe in 6 hurricanes from 15 flights, i.e., 1 flights from Hanna, 3 flights from Isaias, 4 flights from Laura, 3 flights from Sally, 1 flight from Zeta, and 3 flights from Delta. Observations taken within the 500-km radius of the storm center with LWC < 12 g m^{-3} are included. The number of total RSD observations used is 18076 in this study. The detailed description of the data set is provided in Leighton et al. (2022), which shows that the majority of the observations are taken in stratiform environment with relatively weak vertical motion. A closer look at individual storm reveals that most of convective environments are present in Hurricane Isaias (Fig. S1 in the Supporting Information).

3 Methodology

The moment of a raindrop size distribution is defined as:

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

Where m is the number of moments, $N(D)$ is the raindrop size distribution as the function of diameter D. Inserting Eq. 1 into Eq.2 above, we arrive at

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

Given the special property of gamma function,

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

any group of three consecutive moments gives a set of unique solutions of the three parameters for the gamma distribution. After manipulating Eq.3 for three consecutive moments (e.g. m, m+1 and m+2), we obtain the solutions as following,

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

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

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

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

In this study, we explore five different combinations of three consecutive moments, i.e. moments 0, 1, 2 (m012), moments 1, 2, 3 (m123), moments 2, 3, 4 (m234), moments 3, 4, 5 (m345), and moments 4, 5, 6 (m456). It is worth noting that the moments above are calculated directly from the observed RSDs and the calculated moments might not correspond to the IRVs of the same moments. For example, the 3rd moment calculated above is not the same as LWC since the density of water is not taken into account. The performance of gamma fits is evaluated from two aspects: 1) comparing the IRVs calculated from the fitted gamma distributions and that from the observed RSDs, and 2) comparing RSD shapes by calculating the correlation coefficient between the fitted RSD and the observed RSD. The five IRVs used for evaluating the performance of the fitted RSDs are total number of concentrations, mass-weighted-diameter, LWC, radar reflectivity and rainfall rate. The calculations of these IRVs are shown in the following from equations 5-9:

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

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

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

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

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

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 of raindrops and follows Best (1950).

4 Results

4.1 Gamma fitting of RSDs

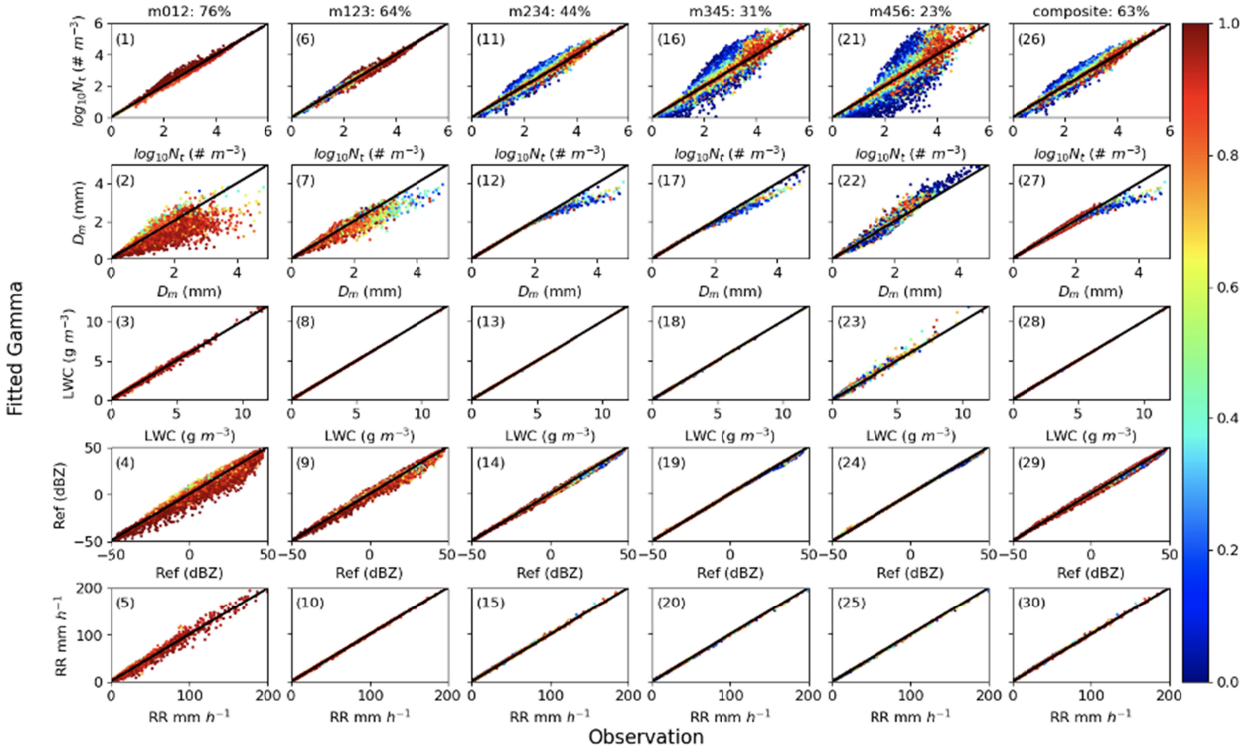


Figure 1: Scatter plots of different IRVs (i.e. total number concentration N_t, mass-weighted-diameter D_m, liquid water content LWC, radar reflectivity Ref, and rainfall rate RR) calculated from the observed RSDs (abscissa) and the fitted gamma distributions (ordinate). The first five columns are from moments 0, 1 and 2 (m012), moments 1, 2, and 3 (m123), moments 2, 3 and 4 (m234), moments 3, 4 and 5 (m345) and moments 4, 5 and 6 (m456), respectively. Last column, termed as composite, are obtained by merging the fitted gamma distribution from all five groups of three consecutive moments such that the best fits of each group are retained (see text for details). The corresponding moments and the percentage of fits that have correlation coefficient >0.9 are shown at the top of each column. The color indicates the correlation coefficient between the fitted gamma distribution and the observed distribution. The black line in

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

Fig. 1 shows the calculated IRVs from gamma fitting using different combination of moments. Each dot denotes each sample of the RSD observations. As shown in Fig. 1, the IRVs calculated from the gamma fitting that uses three lowest moments (m012) have the highest correlation coefficients, and 76% of fits having correlation coefficients >0.9. This ratio decreases with the increasing order of moments and is only 23% for m456. This is consistent with Smith and Kliche (2005) and Smith et al. (2009), who showed that errors of the estimates of the RSD parameters using MM are usually larger when higher-order moments are employed and suggested using the lowest-order sample moments. However, when evaluated by the calculated IRVs, the MM using the lowest-order moments has the worst performance for most of the calculated IRVs. Figure 1 (panel 1-5) shows that m012 slightly overestimates the total number concentration but severely underestimates mass-weighted-diameter, especially for the large drops. Consequently, radar reflectivity is drastically underestimated in m012. The rainfall rate is also underestimated for rainfall rates <100 mm hr^{-1}. The LWC from m012 is in good agreement with the observation. As the order of moments increases to 3, 4 and 5 in Fig 1 (panel 16-20)), the calculation of higher moments reaches the optimum. The LWC, reflectivity and rainfall rate all agree well with the observation. Yet the number concentration degrades significantly, even compared to m234. As the order of moments increases to 4, 5 and 6, mass-weighted-diameter is mostly overestimated and so is the LWC. However, this overestimation is offset by the underestimation of number concentration and consequently both reflectivity and rainfall rate showed good agreement with the observation. This highlights the deficiency of evaluating the fits based on only one IRV. For example, Tokay and Short (1996) showed that calculated rainfall rates from fitted distributions are in excellent agreement with rainfall rates obtained from observed RSDs in their Fig. 1. However, the excellent agreement for rainfall rates alone does not guarantee that fitted distributions represent the observed RSDs well. As seen in Fig. 1, lower order moments produce good agreement with the observations for IRVs such as total number concentration, and higher order moments produce good agreement with the observations for IRVs such as reflectivity and rainfall rate. Middle moments, such as m234, show a good balance that generates overall good agreement for the calculated IRVs, which is consistent with Cao and Zhang (2009). Therefore, our approach is to composite the moment fits from all five groups so that best fits of each group

can be utilized. The composite is compiled according to the following approach. First, the fits from M456 are selected if their correlation coefficient is >0.9 and the error of total number concentration is <10 m^{-3}. The same selection criteria are applied to the remaining samples (total samples minus selected samples from M456), but the fits are selected from M345. Next, for the updated remaining samples (total samples minus the selected samples from M456 and minus the selected samples from M345), the fits from M234 are selected if their correlation coefficient is >0.9. For the new remaining samples (total samples - selected samples from M456 - selected samples from M345 - selected samples from M234), the fits from M123 are selected if their correlation coefficient is >0.9 and the error of D_m is <10%. The same screening process is performed for the remaining samples, but the fits are selected from M012. The rest of the fits come from M234. The overall pattern of this composite MM fitting will resemble M234 but the correlation coefficient is expected to improve. As shown in the last column of Fig. 1 (panel 26-30), the calculated IRVs agree well with the observation in general, as in M234. Yet the ratio of fits that have correlation coefficients >0.9 increases from 44% from M234 to 63% in the composite MM. The distributions of the three parameters for the fits with correlation coefficient >0.9 from the composite MM are shown in Fig. 2.

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

Many studies (e.g. Ulbrich 1983; Zhang et al. 2001, 2003; Brandes et al. 2003, 2004; Vivekanandan et al. 2004; Ulbrich and Atlas 2007) have explored the relationships among the three parameters in the fitted gamma distributions of RSDs. Ulbrich (1983) showed the relationship between N_0 and μ that is deduced from empirical relations between IRVs, such as Z-R relationship, from early studies. Other studies (Zhang et al. 2001, 2003; Brandes et al. 2003, 2004; Vivekanandan et al. 2004; Ulbrich and Atlas 2007, Chang et al. 2009) deduced $\mu - \lambda$ relationships based on fitted RSDs from different data sources. They show that this relationship provides useful information to describe RSDs and improves the accuracy of the retrieved RSDs from polarimetric radar measurements. The relationship between any pair of parameters can also be used in microphysics parameterization schemes. For a one-moment scheme that uses the gamma distribution for RSDs, when one parameter is prescribed, the second parameter can be calculated from the relationship between this pair, and the third parameter can be diagnosed from the prognostic variable LWC and the two known parameters. For a two-moment scheme, two

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

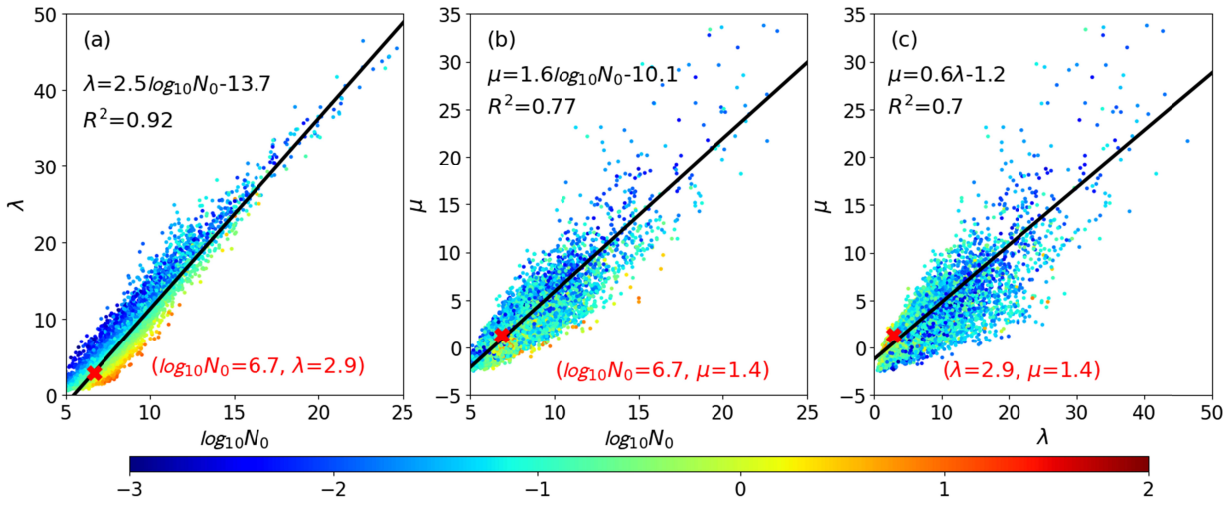


Figure 2: Scatter plots of a) Slope λ vs. $log_{10}N_0$ b) Shape μ vs. $log_{10}N_0$, and c) Shape μ vs. Slope λ. The data points are colored by log_{10} (LWC). The red cross indicates the mode for the pair of parameters and the corresponding mode values are denoted in the red text at the bottom of each 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 trained on a separate bootstrap resampled dataset and using a different subset of the attributes. In this study, we use two attributes, $log_{10}N_0$ and LWC, for the RFs input. Output is slope λ. One common problem in machine learning application is overfitting. In order to objectively evaluate the performance of a machine learning model, the data are always split into training data and testing data. We use 70% of data for training and 30% for testing. The performance of both linear regression and RFs regression is evaluated. As shown in Fig. 3a, R^2 of linear regression model for testing data is 0.91, similar to that in Figure 2a. R^2 of RF regression model for testing data is 0.98, 7% increased on R^2 obtained from the linear regression model. We also split the data by storms, using data from five storms as training data and one storm as testing data, to evaluate if the relationship obtained from the training data can apply to observations obtained from an unseen storm. The results (Fig. S2 in Supporting Information) reveal that the obtained relationships from both linear model and RF regression model can generalize well to the observations from an unseen storm.

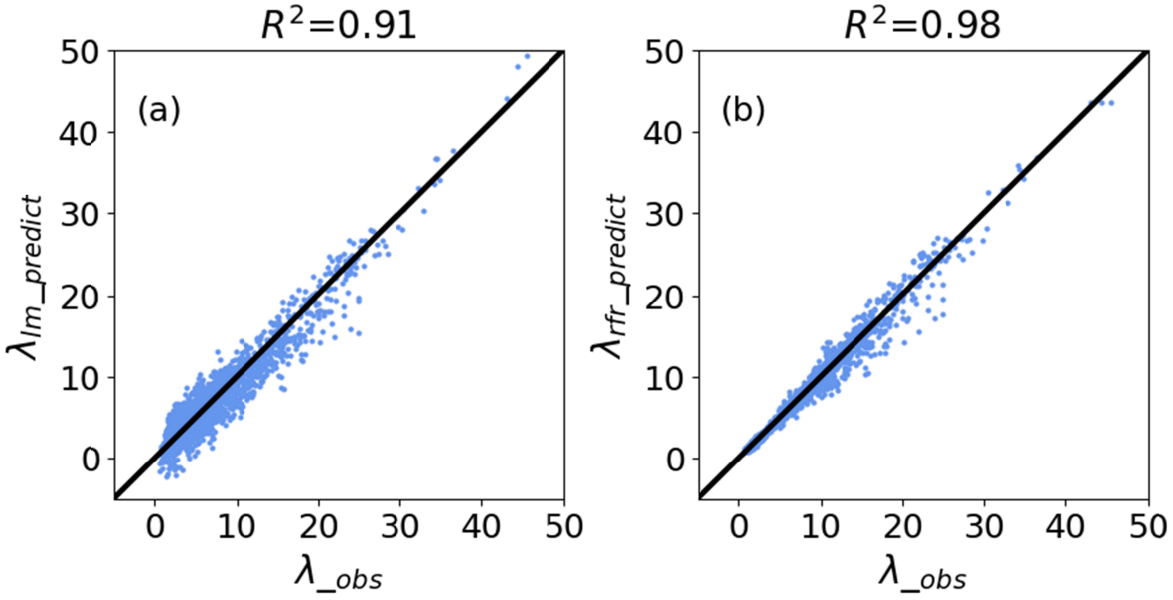


Figure 3: Scatter plots of a) Slope λ from composite MM fitting vs. slope λ predicted from linear regression model (input: $log_{10}N_0$ from composite MM fitting), and b) Slope λ from composite MM fitting vs. slope λ predicted from random forests regression model (input: $log_{10}N_0$ from composite MM fitting and LWC)

6 Discussion and Conclusions

Raindrop Size Distributions collected by PIP from 15 flights through 6 hurricanes during hurricane field program in 2020 are used to study gamma fits in hurricanes. The results from gamma fitting using MM showed that using the lowest orders of moments produces the best fits when evaluated by the correlation coefficient between the fitted and the observed RSDs. Yet, the IRVs, especially radar reflectivity and rainfall rate, are significantly underestimated due to the underestimated mass-weighted-diameter. In contrast, radar reflectivity and rainfall rate calculated from high order MM fits are in excellent agreement with the observations. This excellent agreement is the result of overestimated mass-weighted-diameter and underestimated total number concentration, especially for large drops. The correlation coefficient is much lower for high order MM fits. The central moments, M234, shows overall good performance, yet only 44% of fits represent the observed RSD well, evaluated by the correlation coefficient between the fitted and the observed RSDs. By compiling composite MM fits to extract best fits in each group, the ratio of fits with correlation coefficient larger than 0.9 increased from 44% to 63% without compromising the calculated IRVs.

The distribution of the intercept N_0 and slope λ showed a strong correlation. A linear empirical relationship that is obtained by fitting the entire dataset captures 92% of the variance of the data. The remaining 8% of variance is shown to be closely related to LWC. A RF regression model is able to capture 98% of the variance of the data if inputs include both N_0 and LWC. The distributions of $\mu - N_0$ and $\mu - \lambda$ also show correlation in each pair (Fig. 2b and Fig. 2c) but the scatter is significantly larger than that in $N_0 - \lambda$, making a fitted empirical relationship less representative.

The identified $N_0 - \lambda$ relationship obtained from the RF regression model can not only improve the accuracy of the retrieved RSDs from polarimetric radar measurements by providing useful information to describe RSDs but also reduce the uncertainties and increase the accuracy of bulk microphysics parameterization schemes in numerical models. For a one-moment scheme that uses gamma distribution for RSDs, if N_0 is provided, then λ can be calculated from this $N_0 - \lambda$ relationship with high confidence and the shape parameter μ can be diagnosed from N_0, λ, and the prognostic variable LWC. For a two-moment scheme, the $N_0 - \lambda$ relationship along with two

prognostic variables, LWC and total number concentration, can fully determine the RSD. The accuracy of the microphysics processes in the bulk scheme can therefore be better formulated and so potentially improves the overall performance of the microphysical parameterization schemes.

All the observation data used in this study are from hurricane environment. How well the results from this study can generalize into other weather scenarios needs further investigation using the observation data from various weather scenarios. Nevertheless, the same methodology presented in this study can be adopted.

Acknowledgments

This research was supported by Office of Atmospheric Research Program/Weather Portfolio, Atlantic Oceanic and Meterological Laboratory, and National Weather Service/Office of Science Technology Integration. United Forecast System Research to Operation project of National Oceanic and Atmospheric Administration. We want to acknowledge our colleagues: Drs. Peter Black and Linjiong Zhou, for their insightful comments to improve this manuscript during our internal review.

Data Availability Statement

The authors express thanks to NOAA/HRD Data Support for providing the microphysics observation data (https://www.aoml.noaa.gov/ftp/hrd/data/cloudphysics/2020/)

References

Best, A.C. (1950), Empirical formulae for the terminal velocity of water drops falling through the atmosphere. *Q. J. R. Meteor. Soc.* 76, 302–311.

Brandes, E. A., Zhang, G. & Vivekanandan, J. (2003), An evaluation of a drop distribution–based polarimetric radar rainfall estimator. *J. Appl. Meteor.*, 42, 652–660.

Brandes, E. A., Zhang, G. & Vivekanandan, J. (2004), Drop size distribution retrieval with polarimetric radar: Model and application. *J. Appl. Meteor.*, 43, 461–475.

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

290 Cao, Q., and G. Zhang (2009), Errors in estimating raindrop size distribution parameters
291 employing disdrometer and simulated raindrop spectra, *J. Appl. Meteor.*, 48(3), 406-425.

292 Chang, W.-Y., Wang, T.-C. C., & Lin, P.-L. (2009), Characteristics of the raindrop size
293 distribution and drop shape relation in typhoon systems in the western Pacific from the 2D video
294 disdrometer and NCU C-band polarimetric radar. *J. Atmos. Oceanic Technol.*, 2610, 1973–1993.

295 Gagne II, D. J., McGovern, A., Haupt, S. E., Sobash, R., Williams, J. K. & Xue, M.
296 (2017), Storm-Based Probabilistic Hail Forecasting with Machine Learning Applied to
297 Convection-Allowing Ensembles. *Wea. and Forecasting*, 32, 1819-1840.

298 Gagne II, D. J., A. McGovern, & M. Xue, 2014: Machine learning enhancement of storm-scale
299 ensemble probabilistic quantitative precipitation forecasts. *Wea. Forecasting*, **29**, 1024–1043.

300 Gorgucci, E., and Scarchilli, G. & Chandrasekar, V. (1994), A robust estimator of rainfall rate
301 using differential reflectivity. J. Atmos. Oceanic Technol., 11, 586–592.

302 Haddad, Z. S., Durden, S. L., & Im, E., (1996), Parameterizing the raindrop size distribution. *J.*
303 *Appl. Meteor.*, 35, 3–13.

304 Haddad, Z. S., Short, D.A., Durden, S. L., Im, E., Hensley, S., Grable, M. B., & Black, R. A.
305 (1997), A new parametrization of the rain drop size distribution. *IEEE Transactions on*
306 *Geoscience and Remote Sensing*, 35, 532–539.

307 Herman, G. R., & Schumacher, R. S. (2018), "Dendrology" in numerical weather prediction:
308 What random forests and logistic regression tell us about forecasting extreme precipitation. *Mon.*
309 *Wea. Rev.*, **146**, 1785–1812

310 Kozu, T., & Nakamura, K. (1991), Rainfall parameter estimation from dual-radar measurements
311 combining reflectivity profile and path-integrated attenuation. *J. Atmos. Oceanic Technol.*, 8,
312 259–270.

Leighton, H., Black, R., Zhang, X., & Marks, F. D. (2022). The relationship between reflectivity and rainfall rate from rain size distributions observed in hurricanes. *Geophysical Research Letters*, 49, e2022GL099332. https://doi.org/10.1029/2022GL099332

McGovern, A., Gagne II, D.J., Lagerquist, R., Elmore, K., & Jergensen, G.E. (2019) Making the black box more transparent: Understanding the physical implications of machine learning. *Bulletin of the American Meteorological Society*, 100(11), 2175-2199.

Marshall, J. S., & Palmer, W. M. (1948), The distribution of raindrops with size. *J. Meteor.*, 5, 165-166.

Mueller, E. A., & Sims, A. L. (1966), Radar cross sections from drop size spectra. Tech. Rep. ECOM-00032-F, Contract DA- 28-043 AMC-00032€ Illinois State Water Survey, Urbana,110 pp. [AD-645218].

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

Smith, P. L., & Kliche, D. V. (2005), The Bias in moment estimators for parameters of drop size distribution functions: Sampling from exponential distributions. *J. Appl. Meteor.*, 44, 1195–1205.

Smith, P. L., Kliche, D. V. & Johnson, R. W. (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.

Sulakvelidze, G. K. (1969), Rainstorms and Hail (translated from Russian), Israel Program for Scientific Translations, Jerusalem [available from NTIS].

Tokay, A., & D. A. Short. (1996), Evidence from tropical raindrop spectra of the origin of rain from stratiform versus convective clouds. *J. Appl. Meteor.,* 35, 355–371.

Ulbrich, C. W., (1983), Natural variations in the analytical form of the raindrop size distribution. *J. Climate Appl. Meteor.*, 22, 1764–1775.

339 Ulbrich, C. W., & Atlas, D. (1998) Rainfall microphysics and radar properties: Analysis methods
340 for drop size spectra. *J. Appl. Meteor.*, 37, 912–923.

341 Ulbrich, C. W., & Atlas, D. (2007), Microphysics of raindrop size spectra: Tropical continental
342 and maritime storms. *J. Climate Appl. Meteor.*, 46, 1777–1791.

343 Vivekanandan, J., Zhang, G. & Brandes, E. (2004), Polarimetric radar estimators based on a
344 constrained gamma drop size distribution model. *J. Appl. Meteor.*, 43, 217–230.

345 Zhang, G., Vivekanandan, J., & E. Brandes (2001), A method for estimating rain rate and drop
346 size distribution from polarimetric radar measurements. *IEEE Trans. Geosci. Remote Sens.*, 39,
347 830–841.

348 Zhang, G., Vivekanandan, J., & E. Brandes, Meneghini, R., & T. Kozu, 2003: The shape – slope
349 relation in observed gamma raindrop size distribution: Statistical error or useful information? *J.*
350 *Atmos. Oceanic Technol.*, 20, 1106–1119.

351
352

Figure1.

m012: 76%
m123: 64%
m234: 44%
m345: 31%
m456: 23%
composite: 63%
(1)
(2)
(3)
(4)
(5)
(6)
(7)
(8)
(9)
(10)
(11)
(12)
(13)
(14)
(15)
(16)
(17)
(18)
(19)
(20)
(21)
(22)
(23)
(24)
(25)
(26)
(27)
(28)
(29)
(30)
$log_{10}N_t$ (# m^{-3})
D_m (mm)
LWC (g m^{-3})
Ref (dBZ)
RR mm h^{-1}
Fitted Gamma
Observation

Figure2.

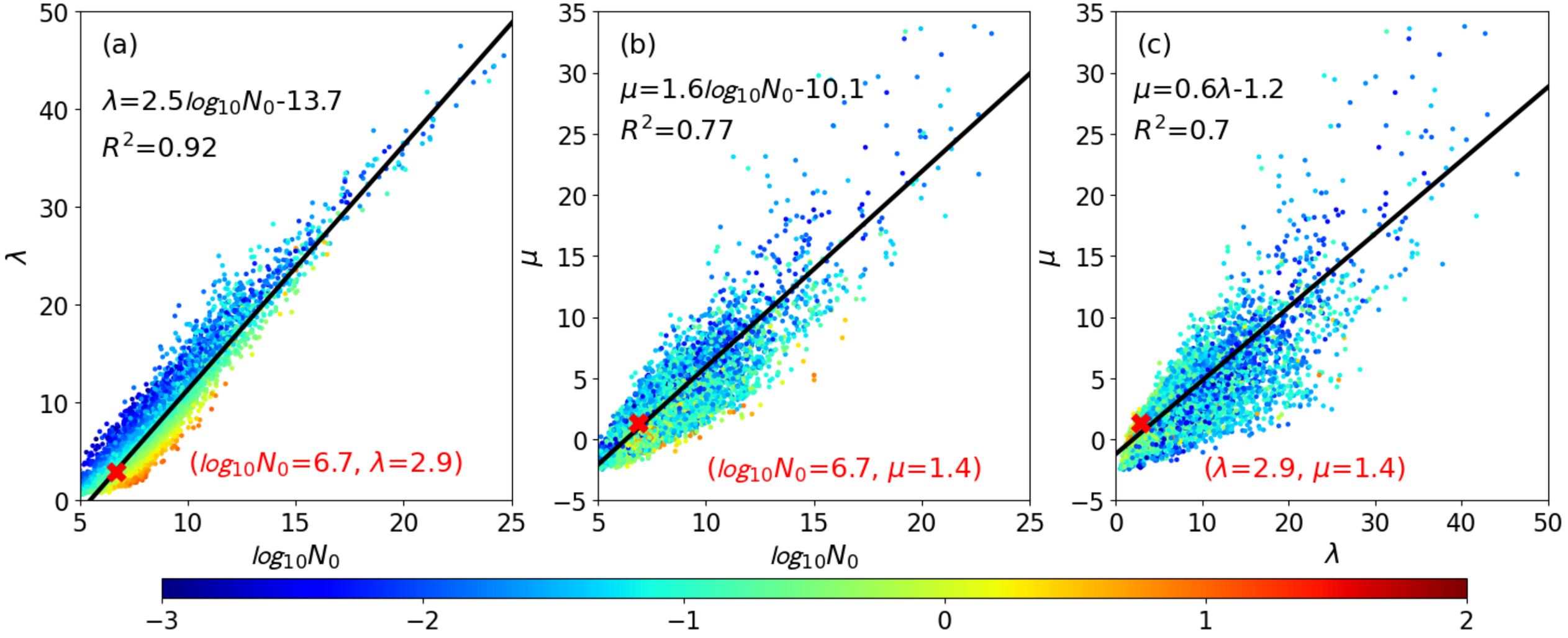
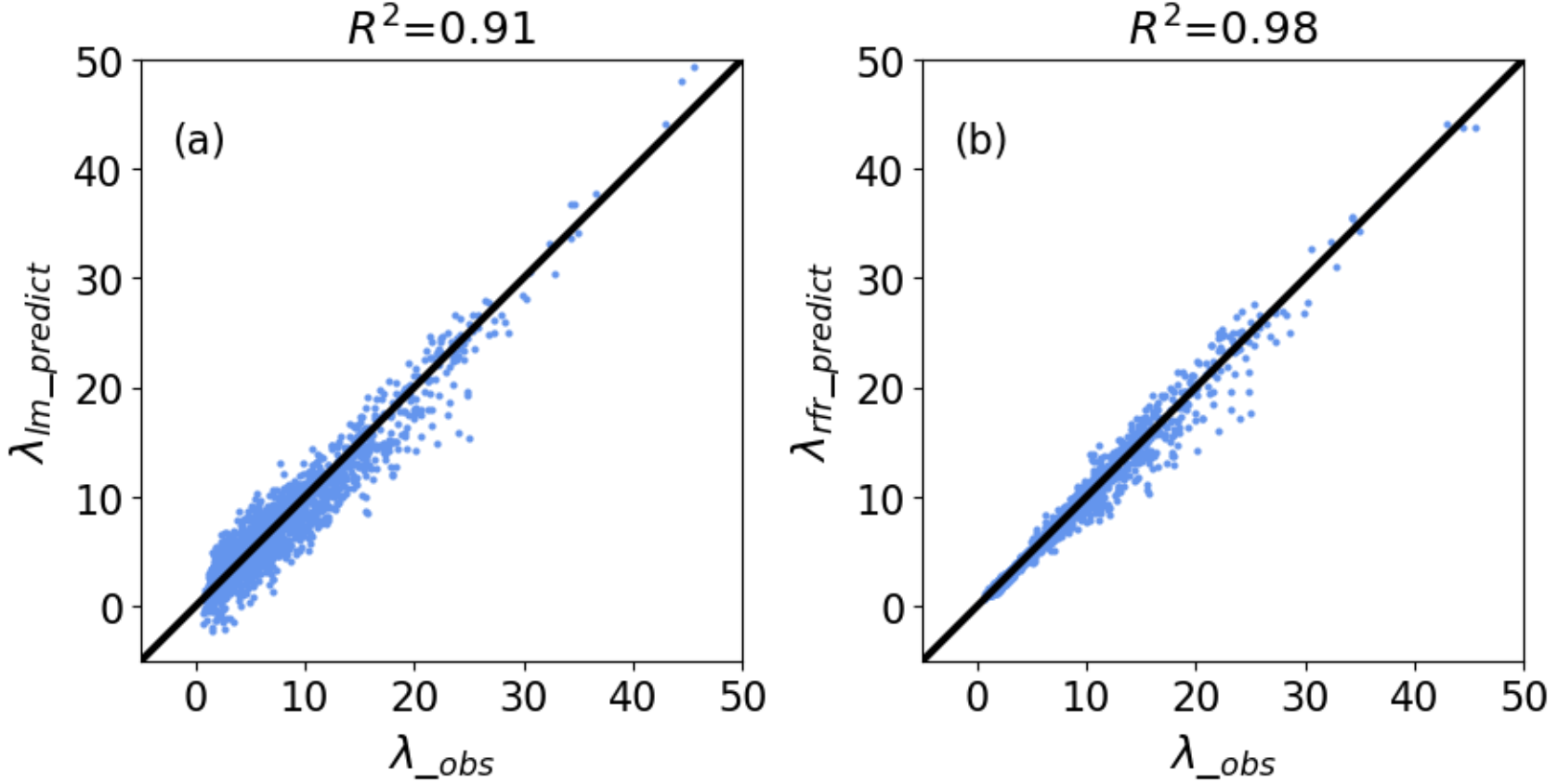

(a)
λ=2.5$log_{10}N_0$-13.7
R^2=0.92
($log_{10}N_0$=6.7, λ=2.9)
λ
$log_{10}N_0$
(b)
μ=1.6$log_{10}N_0$-10.1
R^2=0.77
($log_{10}N_0$=6.7, μ=1.4)
μ
$log_{10}N_0$
(c)
μ=0.6λ-1.2
R^2=0.7
(λ=2.9, μ=1.4)
μ
λ
−3
−2
−1
0
1
2

Figure3.


R^2=0.91
(a)
$\lambda_{lm\_predict}$
$\lambda_{\_obs}$
0
10
20
30
40
50
R^2=0.98
(b)
$\lambda_{rfr\_predict}$
$\lambda_{\_obs}$