Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2019-34-RC2, 2019

Atmospheric Chemistry and Physics Discussions

Interactive comment on "On the distinctiveness of oceanic raindrop regimes" *by* David Ian Duncan et al.

Anonymous Referee #2

Received and published: 3 March 2019

The paper presents many interesting results using a wide variety of relevant and novel data sources. The aim and motivation of the paper are well-justified - the paper addresses the need to physically understand rain variability from observations and to accurately retrieve this information with the currently-available suite of satellite measurements. I am highly supportive of the paper. However I find many areas that need improvement because the way things are written currently (vaguely, with results confused with discussion) leave many statements open to interpretation.

The subject matter of this paper is worthy of publishing, but the paper first needs to be revised in several ways:

1. Reorganize to separate Data Information / Tools (the datasets, their intricacies, and the models you ran them through) from Results (what your methods produced)

from Discussion (as a logical conclusion of the results shown, what do the results mean?; what are the implication of results based on the methods you used?). Right now, methods, results, and discussion are mixed throughout each paragraph. This often means the authors often repeat themselves. The main objectives of the paper and main results of the paper are obscured by this mixture of results + data details + discussion in each paragraph, and from paragraph-to-paragraph. There is not enough logical order within each paragraph or from paragraph to paragraph. Right now, the paper is separated thematically by the data used, which also serves a purpose but is tiresome for the reader. As written currently, the authors don't leave a well-defined trail for the reader to follow, instead they loop back and forth over their tracks several times and lead the reader to wonder where the story is going, where has it been already, and why. Consider implementing a roadmap of Background, Motivation, Objectives, Data, Methods, Results (just the facts - just state exactly what you did and present the figures objectively), Discussion (interpret the results, draw nuanced conclusions, and comment on the significance/implication of the results for other applications), and Summary (tell us how you met the challenge stated, hopefully clearly, in the Motivation/Objectives. Progress from section to section, don't repeat yourself throughout.

a. The most egregious instances of mixing discussion with results are in Sec. 3. The beginnings of many of these paragraphs should be deleted b/c they are redundant, or the methods are restated. The last sentences of many of these paragraphs appear to opinions of the authors and should be separated from the results and kept in a discussion section. Ex - last sentence of 3.1 is not a logical conclusion from data shown, it's based on the experience of the authors and their interpretation.

2. Rewrite/Revise each sentence for literal truth. Each statement should be made to be physically true, word for word. So many of the sentences in this paper contain vague, non-specific, perhaps physically impossible language that is abstract and doesn't really teach the reader anything. The authors need to more specific detail in several places so that their results are digestible, understandable, and reproducible. In many cases,

I did not understand what the authors even did in terms of methods, how they drew the conclusions stated, what certain plots really showed, or how to draw conclusions from these plots. In many cases, problems/issues are alluded to (e.g. uncertainty, constraints, sensitivity), but the specifics or causes of these problems are withheld so the reader is left with suspense - what could the problems really be and why are they there? We don't know, because the authors allude to issues without explaining them. Examples of vague language that doesn't really explain anything physical:

a. GMM easily generalized to a wide variety of data distributions * What is meant by easy? What does generalize mean, physically? How wide? Of what variety? What are data distributions, in this context?

b. Underpinning OceanRAIN is the ODM-470 optical disdrometer * Does a disdrometer literally underpin (lie beneath, support, substitute for weaker material) the data network? No. These instruments are used to collected the data amassed in this network.

c. Limited sensitivity of DPR to small drops * How sensitive? How small of drops can be seen?

d. Has a more bimodal distribution * How do you quantify what is more or less bimodal? The PDF either exhibits two peaks in frequency of occurrence, or not.

e. Commonly occuring, common forms, commonly used DSD, and other uses of common * Used by whom? How many times? What makes it common? How can a DSD be commonly used? All uses of common should be replaced by a descriptor that is physically-founded and/or citable, as in which studies/models used it, or how many?

f. Data that contain discontinuities between size bins and some degree of instrument error * What discontinuities? How is this quantified? * What degree of instrument error? And why?

g. Unmolested observations * NO - delete/revise for something more physical and less

politically/emotionally charged. h. Magic bullet

* NO - delete/revise for something more physical and less politically/emotionally charged. i. “Robust”, or even worse “very robust”

* Search all instances of this word and replace with a physically-based and specific metric. Error, uncertainty, skill, accuracy... what do you mean by robust? How is that quantified or assessed?

The paper could be shortened, simplified, and made to be more digestible and impactful if the sentences were simplified. The authors should carefully edit the manuscript to just state the facts, use simple sentence structure, and use specific numerical details (above 2.5 mm, greater) instead of subjective, nuanced, literary adjectives or descriptors. Examples of overly verbose, passive voice, unspecific, and incomprehensible statements often contained in this manuscript: j. OceanRAIN observes a wider distribution of Dm that is “most noticeably distinct from GPM results for small drops”. * You lost me with the phrase in quotes... what does this mean? How small? How do you quantify what is “most noticeable distinct”? Are there more physical or statistically significant ways you could try to prove your point here?

k. Functional forms, models, fits... * Pick one and stick to it? Are they different?

l. Underconstrained, unconstrained * By what metrics? How is this defined? It's mentioned several times that retrievals are underconstrained/unconstrained, and this is identified as an issue. However, it's not clear what the issue really is. What is missing from the constraints?

m. Have the correct RWC * Correct to: conserve total LWC compared to observations

n. not particularly Gaussian * Particularly? It is either is or is not Gaussian

o. This can then be compared between DSD representations * I have no idea what this means

p. Warmer ocean surfaces witness greater densities of drops * Since oceans do not have eyes or consciousness, it might simply be the case that rain exhibits higher number concentrations when it develops and falls over water with warmer SST.

q. May or may not make much difference. * Compared to what? What differences are you referring to? What would change, or not?

r. Bulk radiative, bulk properties, bulk parameters * What is meant by bulk here? What does it summarize? It seems like a filler word that misses the point, if there is one. Vertically integrated? Averaged? From 1-m depth, over the lower km? Use physical descriptions instead of vague adjectives that are open to interpretation.

s. Peaked distribution, smaller drops, fewer drops, higher, lower, more, less, tail, maxima, very weak, very strong, large range, lower resolution, relatively similar, overall, generally, overall behavior, extreme values, less spread, slightly less pronounced, scant, much of, underrepresentation, markedly different, not particularly, not especially distinct, and any instance of slight, almost, barely * All of these extraneous, subjective, hand-wavy descriptions need to be made more physical/numerical or deleted: Mode = most frequent value, was it the actual maximum value? How much higher? How small? What diameter ranges? How extreme? How low? Very is a filler word. Just quantify what you mean. The data should speak for themselves. Overall and generally are also meaningless filler words. Quantify your analysis to prove your point. Use numbers to justify your interpretation. * What does underrepresentation or overrepresentation mean? See discussion of Fig. 3 on Page 9-10.

t. Posterior * Replace all instances of this with something more plain-language or physical. Not immediately apparent how an after the fact (?) step is executed.

u. Represented by this formalism * ?

3. I'm unsure about the title after reading the paper. The authors conclude that there are several forms of DSD shape/solution, all of which overlap a bit, and their large

datasets include instances of DSD that oscillate around these forms (as shown by the machine learning). Then they make conclusions about how well in situ observations of DSDs are represented with current satellite DSD retrievals (sensitivity problems appear to limit the capabilities of the satellites), and how much scatter should exist in satellite retrieved Zh and Tb given the observed DSD (quite a bit, and it varies based on frequency). The paper really doesn't separate out "raindrop regimes"... which would indicate to me to be physical distinctions like monsoon phases, convective v. stratiform, different meteorological conditions, etc. Perhaps the impact and message of the paper is best summarized as: "On the distinctiveness of oceanic raindrop distributions in observations and satellite measurements"

4. In intro: "A lack of globally representative DSD"... do we lack that? What is lacking, specifically? Are you saying that we need to find a single DSD that represents all DSD across the globe (implied by the way this is written), or that we need a representative dataset? Unclear. 5. The authors don't give enough credit to the oceanic-ish DSD measurements that are still being reported at Manus Island and Kwajalein Atoll. These are not representative of other DSD in other regions, but they do contribute something important; not all the DSD studies have been confined to land or coast (Thompson et al. 2015 and 2018, and this data is also used in Dolan et al. 2018).

6. The authors make up a quantity of RWC and RWP, when it seems that literature to-date uses LWC and LWP to mean the same thing. Paper should conform to these precedents instead of inventing or using their own terminology. It's clear in the paper that only rain cases are used, not ice- or mixed-phase precip, at least at the ground.

7. Fig. 2 - each of the plots must be spaced farther apart in the Latitude dimension so that the distributions can be viewed as distinct. The plots run into each other and cannot be distinguished one from the other. Therefore, the point of this plot is lost. Also, It's not clear what the latitude averaging bin width was here. It's centered 20deg apart - is that the center value +/- 10 deg in each direction? Explain in text and caption

8. The units of Nw need to be checked - it's often discussed in the text as Nw when it's really logNw that is plotted. The parentheses should be adjusted to explain the units properly. log10(Nw [mm-1 mm-3]). Log10Nw doesn't have the same units as Nw itself.

9. Each of the box and whisker plots need a legend that explain the box width, lines dots, cross hairs (what are the percentiles, etc.). It's too much work to read the caption and try to interpret the plot at the same time.

10. Add a plot of the map of observations used, so that your Fig. 1 and comparisons between latitude bands can be better understood. The text mentions that more or less data was collected in certain regions, but the reader can't deduce that because no plot of the data extent is shown.

11. Fig. 4 is just about incomprehensible. The conclusions about this figure do not appear to stem logically from the plot, and the plot itself is confusing. If you are making a comparison between models and obs, that is one thing and can be shown a certain way. Then if you also want to compare latitude bands, you should show the PDFs from those latitude bands first before you plot the difference. The difference here doesn't really make sense either because it doesn't seem physical to subtract a tropical PDF from a higher lat PDF; the result doesn't really have any physical significance (or at least the significance is not well-explained or logical). Suggest deleting Fig. 4 and showing the real tropical and high latitude PDFs. Then see if your stated conclusions are supported by the real data. The entire discussion of Fig. 3-4 needs to be carefully curated to be sure that the interpretations of the authors are readily apparent from the plots shown; I couldn't really see how they made all the conclusions that were stated. Particularly the regional dependence, but also the under vs. over "representation" issue. Fig. 4 needs titles to distinguish them and explain more about what is shown.

12. I kept forgetting what GMM and MGD meant. . . these acronyms are used throughout and their physical significance became lost several pages in. Suggest using plain language. See. Fig. 5 - it's not clear what the significance of the acronyms are for the

purposes of showing the data.

13. Typesetting should be improved a. The lack of line numbers on each line makes it difficult for the reviewer to suggest particular comments and corrections in each case. The authors should number each line in the future. b. The paragraphs would also be easier to distinguish if their indentation was larger. It kind of all runs together c. The references within parentheses are wrong. For example this is typed often: (also used by Duncan et al. (2019)) < see double parentheses at end here. d. Most journals forbid sentences to begin with the name of the variable and will insert words to rectify this situation, whether or not it makes physical sense. The authors are encouraged to fix this the way they want before it gets changed for them (see paragraph above equations on page 2). Completing these sentences also will help ground them in concrete, physical terms: e.g. correcting to be: "The values of N0 and NW. . ."

14. Equation (3) is really 3 equations. . . separate them?

15. Nw is normalized by LWC. make this clear in its discussion on page 2 and elsewhere.

16. Petkovic et al. 2018 reference page 2 - this is not the definitive nor the first paper to study these things. Reference prior work more appropriately to justify this statement - other canonical papers?

17. References are made to OceanRAIN-M . . . what is the -M??? Not defined.

18. The ODM470 doesn't sense drops smaller than 0.4 mm because the voltages are too noisy below this inferred level (the sensitivity is incorrectly stated as 0.3 in the manuscript). From Klepp 2018: "The first 12 bins ranging from 0.04 to 0.36âĂĹmm in size are not recorded because they are prone to contain artificial signals caused by ship vibration." This is a significant downside to the instrument that needs to be explicity stated in the manuscript. A large portion of the DSD spectrum is contained below this level (Thompson et al. 2015, Thurai et al. 2018). Also, the ODM assumes a

fall velocity and does a vector mean of the fall velocity + wind through sensor. This is an approximation and, in my interpretation of the ODM methods, a potential source of error in the computed number density.

19. “Is an issue faced by all disdrometers” -> ya, but it’s the worst for ODM compared to other disdrometers currently used. See comment above.

20. Minimize effects of turbulence -> minimize under-catchment of drops and minimize the effect of drop splashing against the sensor from multiple directions. Turbulence will still exist. It’s just that the orientation of the instrument attempts to catch as many drops as it can.

21. Why would the disdrometer NOT “show no difference in accuracy between oceanic and continental cases”? Why would differences occur? It should work the same, except that while at sea, it has to reorient into the wind and a vector mean calculation is done to assume fall velocity+wind through the sensor, which could be susceptible to flow distortion.

22. Introduction of the “Combined Data” is confusing. The heading of that subsection is also confusing. What does it mean?

23. State the different vertical resolutions of the two DPR frequencies. It’s mentioned that they differ, but not what they actually are.

24. The specifications of the GMI and DPR (geometry, resolution) are scattered about the last paragraph of 2.2. Make it more concise.

25. “Ground-based data from Ocean-RAIN-M” —— Ocean rain is over ocean! Not land.

26. How is clutter classified in satellite retrievals? Sec. 2.2. Seems like it would impact your analysis.

27. Provide citation for GMM upfront at its first introduction

28. 90,000 points seems REALLY low compared to the size and length of time of the

OceanRAIN dataset. Is this correct? I’ve collected ODM data over a month and gotten more than 40,000 points of usable data.

29. What are the size ranges of the 60 size bins? And what are their approximate spacing?

30. In several places, the authors assert that assumptions and fits haven’t been used on DSD obs, but then they say that the data are normalized and that a “nominal shape parameter is assumed” . . .. This seems contradictory. And the authors fail to mention how the data are normalized - by liquid water content? Similar to Nw? how was the parameter chose?

31. Provide citation of BIC when introduced

32. The authors make claims that they are testing for differences in DSD in different regions or locations, but really they only separated data into very wide latitude bands (Fig. 1, Fig. 2). This distinction, and the implication of the real results, need to be me explicit so as not to overstate the significance or conclusions. Search all uses of “region” and “location” to determine whether you actually mean “latitude”

33. Similarly, the authors make mention of a stratocumulus region, but have not proven where this is or justified their interpretation of that based on a map of the data - which should be added to justify statements made throughout the analysis in Sec. 3

34. The differences in DSD based on SST seem readily explainable from the Clausius Clapeyron equation - warmer SST leads to higher saturation vapor pressure of air, so moisture content can be higher. – see Fig. 1 and Fig. 2. Nw is directly proportional to LWC (your eq. 3), so result of Fig. 1 lower right is awesome but perhaps not surprising.

35. “Blue” dots in Fig. 2 don’t show up. Looks black.

36. What are the bin sizes of latitude and SST used to create this plots?

37. Captions and discussions of FIg. 2 say Nw but you plotted log10Nw

38. The satellite is not sensitive to smallest drops, and probably also not to the lowest number concentrations. This seems to explain some of the differences in flat vs. peaked DSD shape in FIg. 2.

39. Evaporation below the lowest altitude of GPM would impact the smallest drops first (see Matthew Kumjian et al. papers with observations and models). So, this effect, if it's occurring, would eliminate the small drop portion of the spectrum first, most likely. Since small drops usually present in large number concentrations (Thompson et al. 2015, Dolan et al. 2018), this might mean that the satellite also misses large number concentration DSDs as a result of missing the smallest drops.

40. The exponential DSD is the most commonly occuring, which is why large combined or averaged datasets of DSD often exhibit a shape of this kind (explained in Bringi and Chandrasekar 2001). However, this book also states why you wouldn't expect high-frequency observations of DSD (such as from 1-min observations) to look exponential. The Marshall Palmer was also based on stratiform, steady, UK rain… so again, it's most representative of this steady-state, averaged, stratiform, weakly-forced rain DSD.

41. It is stated on page 10 that rain rate is 3rd moment of DSD; it's actually 3.67th moment. LWC is 3rd moment (Bringi and Chandrasekar 2001).

42. "Result in a small overestimation of rain rates by 0.06 mm/hr or 1.9%" - unclear how this was calculated? At all rain rates? Or overall error? I'm confused how the error in mm/hr can be just one number?

43. Raw observations - it's not really raw. You selected certain sizes, the data have been converted from voltages at fractions of a second to DSD parameters based on A LOT of assumptions contained in the Klepp papers. They are not raw, but they are observations you described already in methods. Just don't use a new qualifier to describe them a different way.

44. How were the "random samples" and "random sampling" and "randomly sampled

subsets" performed? Chunks of data taken at random? How much? Explain.

45. Page 11: "curve was chosen"... I just see lots of curves on Fig. 5. How was it chosen?

46. Last sentence of 4.1 - was this also seen by Dolan et al. 2018?

47. Text at top of Page 12: it seems that you are trying to say that all GMM output DSD fits depart from the exponential form, so that observed mu almost never equals zero.

48. FIg. 6 - label legend better in terms of GMM and GMD parameters (and explain in plain language, it's not clear here). In caption, differentiate solid and dashed lines. Water content should be LWC. Keep consistent terminology/abbreviation throughout.

49. Page 13: rare to have -> rare to observe; observations -> minutes

50. Last few sentences before Fig. 7 does not make any sense; overly literary and wordy. And it's not clear how these conclusions can be drawn from what is shown.

51. It's not clear how interpretation of BIC in Fig. 7 leads to the conclusion made by the authors. I'm totally confused how there can be a singular set… seems like an oxymoron and I don't know what this is intended to mean. It actually appears like 2 modes describe most of the variability (highest BIC) and then Ngmm from 3-6 does fairly and equally well.

52. The last paragraph of Sec 4 is very problematic and hard to follow, and hard to deduce the conclusion from the results shown. The claims appear unsubstantiated as written. I don't know what "decoupled from RWC" . . . there was no regional comparison performed… use of "not particularly" and " not especially distinct" are handwavy and should be replaced by quantifiable and specific terms.

53. What are "variational" systems? Page 14

54. "Water content" is stated… but else where its RWC. stick with consistent terminology/abbreviation once defined, otherwise it sounds like you are introducing a new

concept or field.

55. Page 14: "was used to perform model simulations" . . . of what? Incomplete sentence.

56. Page 14: what are the custom distributions, or habits? Mentioned but not explained sufficiently.

57. Unclear what g m-3 cloud looks like or what this number is supposed to represent.

58. Define TESSEM2

59. How was the change in TB computed? From level to level or from certain simulations? The description of Fig. 8 and these simulations is very confusing. I don't really know what you are trying to do or how you are doing it. Clarify. WHat are the units of TB here? What about units of RWP? Units need to be stated in the text. I don't understand what impacts are being tested here , "The 89 GHz shows little impact" . . . impact of what on what?

60. What do you mean by"largely cancels out its emission signal?" Page 14

61. What are the sizes of cloud droplets and rain droplets that produce these stated differences in emission vs. scattering?

62. Fig 8 and Fig. 9 and other figures where error bars (of some kind?) are shown - need to define on legend what the errors bars and whiskers and dots really mean.

63. Add titles to Fig 8 so that you can tell from looking at the plot what the differences between a and b are.

64. Fig. 8 discussion: "net response" should be "mean value" according to your plot. Be specific about what you are discussing. I'm confused what "response" means in this context.

65. The averaging window of 6 min seems too short to prove your point. Assume

that ships transit at 10 kt (nautical miles per hour) - a good assumption, or you can test it with all the meta data in the ODM dataset. Anyway, figure out how long in time you need to average in order to approximate a 5 km or degree or whatever distance spanned by a single pixel of satellite data

66. "Should not be markedly different." . . .. If what? Compared to what? A comparison or warning seems to be made, but it's not clear what it is.

67. "Non-raining points" – do you mean that you did not average zero values? Unclear

68. Fig. 9 - why aren't data shown above 40 mm/h? Surely they exist in ODM?

69. The author's conclusion that the error is way bigger for the blue line is not well-supported. The errors are only different in blue v. red at the very highest rain rates (above 10 mm/hr). Consider revising interpretations for clarity / consistency.

70. Page 16 "largely been predicated on limited land-based observations in the past" - including some island observations from Manus, Gan, and Kwajalein.

71. Page 16: here it is stated that DSD have limited dependence exists on SST, but in presentation of Fig. 1 you remarked that there appeared to be an SST dependence. Inconsistent.

72. Page 16 in several places: datasets "observe" -> exhibit. Not literally true.

73. Summary and Conclusions: your latitude bands are very wide and may inhibit your ability to draw conclusions about DSD variability that could be present in regimes such as the monsoon, ITCZ and SPCZ, subtropical highs, western boundary currents that are really warm, cold current regions, etc.

74. "Function to encapsulate" and "appears more applicable" – > unclear. Revise.

75. Line 3 page 17: spatial and temporal considerations

76. Unclear what this sentence means "Its use can cause systematic. . ..

77. Sentence with “for about 3% of cases . . .” = run-on sentence. Revise to simplify it.

78. Page 17 lines 12-24 make very little sense. Revise for clarity a. Uncertainty = standard deviation? Width of distribution? means? b. What is “radiative uncertainty”? c. How “rapidly” d. What is meant by low vs. high frequencies? e. “Half the radiative signal” - do you mean the mean? f. What is “true for passive and active simulations” ?? uncler based on what has been shown or discussed g. Unclear why “retrieving the DSD would shrink these ranges” unsupported from what is shown or explained, or maybe it just doesn’t make sense h. Lines 19-24 don’t really make any sense. Revise. Unclear. i. It’s confusing that you bring up dewpoint temp and RWC on line 28 because you didn’t discuss Td anwhwere else but you did use RWC. j. What is meant by “limited spatiotemporal sampling of OceanRAIN”? – Please include a map in the revised manuscript k. What references can you provide to justify why DSDs are more uniform over land. They seem to exhibit more modes over land because hail and mixed-phase and very intense convective microphysical processes can occur. Unclear from what is written.

I tried to understand what the authors did, which is actually a lot of work and very interesting.

The paper synthesizes a wide range of data types:

This paper presents DSD data from 90,000 individual minutes of rain sampled throughout oceans. The data were collected with the ODM-470 disdrometer mounted on 9 research ships between 2010 to some time in the present (the cut off date of this analysis was not specifically stated). The study also makes use of satellite measurements and derived quantities at each of the times and approximate locations of the in-situ DSD data. The satellite data used in this study include the vertical profile of DSD parameters: Nw (lower - how low?? - vertical resolution) and Dm (250 m vertical resolution), while the shape parameter mu is assumed to equal 2. These DSD parameters are retrieved along the DPR vertical profile of . . . [what DPR values are of interest??

reflectivity, Zh? The retrieved field of interest from DPR is not specifically stated but should be]... at 5 km horizontal resolution at two frequencies (Ka, Ku), and also the GMI brightness temperature, Tb, at a [wider?? Not specifically stated??] horizontal dimension at several frequencies (10-190 GHz). This set of data is from the “GPM Combined” dataset.

The analysis of these data is two-fold.

1) For both the in-situ and satellite data of DSD, compare these observations to functional fits and shapes derived from (a) numerical solutions to physical equations (b) machine learning models. 2) Using only the in-situ data, investigate how the DSD and the DSD’s liquid water content/path impact two remote sensing quantities (brightness temp and radar reflectivity). The mean dependence of the satellite quantities on these DSD properties is assessed, as well as the variance and standard deviation for different bins of LWC/LWP/drop size. This task is performed at different frequencies used operationally in satellite rain retrievals passive microwave and radar systems. By examining different frequencies, the authors are able to compare the sensitivity/variability of satellite measurements to certain DSD properties as a function of satellite frequency. The standard deviation/variance in modeled satellite retrievals resulting from initialization with the full dataset of DSD is assessed, and referred to as the forward model error.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-34, 2019.