Review of “The Virga-Sniffer – a new tool to identify precipitation evaporation using ground-based remote-sensing observations” by Kalesse-Los et al.

The manuscript describes the Virga-Sniffer, which is a python-based open-source tool to detect precipitation that does not reach the ground, i.e. virga, on the basis of cloud radar, ceilometer and optional additional measurements. The algorithm was developed and tested using ship-based measurements carried out on the RV Meteor during the EUREC^{4}A campaign. The virga-detection is evaluated against the Cloudnet target classification algorithm with good agreement. Additionally, statistics of virga occurrence as well as the relationship between virga and cloud macrophysical properties (namely cloud base height and cloud depth) are presented for the EUREC^{4}A campaign.

The Virga-Sniffer presented seems to be a well done tool to aid studies of precipitation evaporation. It uses rather commonly available measurements, which together with the modularity and configurability of the tool likely make it applicable beyond the data set used by the authors. I especially appreciate the critical discussion of the limitations of the virga detection presented by the authors. However, although the content of the paper is fine, I find that the presentation at parts is lacking. Especially the description of the Virga-Sniffer is hard to follow, and therefore the manuscript needs improvement before it can be accepted for publication. Please find below my detailed comments, first regarding the description of the Virga-Sniffer, followed by further minor comments.

Comments on the description of the Virga-Sniffer (Sections 3 – 3.3)

1. It is not clear what processing is optional and which steps are always performed, and which of the description applies specifically to the processing of the EUREC^{4}A data set. The authors might need to make a more clear separation of the general description of the algorithm and the EUREC^{4}A specific processing. The authors should also check that the optionality of different steps is clear and uniformly presented across the manuscript.

2. It is also not clear which of the many thresholds given are user-configurable, and I kindly ask the authors to clarify whether some thresholds (if any) can not be chosen by the user.

3. There seems to be discrepancies between the text and Fig. 2, which shows the workflow of the algorithm. I first got the impression that Sections 3.1-3.3 correspond to the three orange boxes in Fig. 2, each section describing one box, however parts of the text describe processing that is shown in a different orange box. Furthermore, there is processing described in the text that is not included in the figure, and elements in the figure that are not included in the text, as far as I can tell. Specifically, I am missing the description of the *Range-gate mapping* (orange box 2), the *smoothing* that is presented in the ellipse below orange box 1, and the *Count valid data* (orange box 3) in the text. Could the authors add the description of these algorithms in the text, or make it more clear where a certain algorithm description is related to the corresponding element in Fig. 2? I urge the authors to check that the Fig. 2 and text logically relate to each other, and suggest the authors use Fig. 2 more to guide the reader through the multiple processing steps.

4. P. 7 L. 140-148. The description of the overall structure of the VirgaSniffer could be extended. I believe providing some more top level description of the processing chain would be helpful to understand the following sections and the context in the processing chain that these occur in. Although the three parts of the virga detection (somehow related to the 3 orange boxes in

Fig. 2) are introduced, introducing also what happens outside these boxes, and where in the manuscript these different parts are described, would be helpful for the reader.

5. Figure 2. There seems to some parts of the flow missing, i.e. some arrows only lead to somewhere but don't start from anywhere, and there is a lonely ellipse "smoothing" that has no input put feeds into several polygons or ellipses. It also strikes me somewhat odd that from the orange box 2 *Precip. & cloud detection* there is no arrow leading to the polygons CBH and CTH in the *Output dataset*. Could the authors check that following the arrows in the figure one can indeed trace the data processing chain, and update the figure where needed.

6. Figure 2, Orange box 3. There are some options shown in the figure, (e.g. mask_clutter=True, mask_vel=True). Are these default options, or the ones used for the EUREC^{4}A data set? I kindly ask the authors to add this information in the figure caption.

7. P. 7 L. 150-153**.** I find the introduction to this section confusing, and it is hard to keep track of the different configurable and non-configurable processing and in which order things are done. To make it easier to follow, I suggest changing the order so that the smoothing that is done as a first step, and which I gathered to be non-optional (however I'm not sure) would be introduced first. Following this, the optional, user-configurable modules could be introduced. Another option would be to first simply introduce the five modules, and in a separate paragraph explain how they are used.

8. P. 7 L. 152. "used settings and thresholds are…" could the authors specify where the settings and thresholds are used, as default values? For the EUREC^{4}A data set?

9. P. 7 L. 154-162. I cannot follow how this processing is done. What are CBH layers, and how is a data coverage threshold or a mean value for these calculated? As far as I understand, for each ceilometer profile a number of cloud base heights are detected, which I would expect to be related to different cloud layers, and thus I don't understand how e.g. a mean value would be calculated for a cloud layer for the one data point available. Is perhaps some kind of time window investigated? I also have trouble understanding the logic of the split and merge modules. Perhaps the authors could consider adding an illustrative figure to help the reader to follow their reasoning.

10. P. 8 L. 163. Also here I don't understand how it is possible to have nan values for the lowest CBH layer, I would expect the lowest CBH to have the value given by the ceilometer for the first cloud base height, or clear sky conditions. Perhaps the clarification of the definition of a CBH layer makes also this more understandable.

11. P. 8 L. 163-164. Is the running-median filter applied to the LCL data before or after replacing the lowest CBH balues with the LCL?

12. P. 7-8 L. 154-165 and Figure 2. The CPH preprocessing modules have different numbering in the figure and in the text (0-4 in the figure, 1-5 in the text). In the appendix (P. 22, L. 430) the 0-4 numbering seems to be in use. I suggest uniform notation to avoid confusion.

13. P. 8. L. 165. What is done at this step? What is the parameter being smoothed? Clarification needed.

14. P. 9 L. 166-168. From this description I was not sure how to the EUREC^{4}A data set was processed. To avoid ambiguity, perhaps the authors could here also give a list of the modules in order of at which the processing was done.

15. P. 9 L. 171-172. Is the linear interpolation described here optional or not? Also, the interpolation is not mentioned in Fig. 2.

16. Is the ceilometer CBH data brought to the same temporal and vertical resolution as the radar data? If not, how are differences in the temporal and vertical resolutions dealt with?

17. P. 9 L. 182. Here the authors argue that a gap of 700 m, which is used as a threshold to detect precipitation associated with a cloud base height, is small enough to not mask out any lower cloud layer, however, later they show that it can happen and the authors discuss the difficulty on setting this threshold. I would find it appropriate to use less definitive language here, and perhaps write that the difficulty in setting this parameter is discussed later in Sect. 5.4.

18. P. 9 L. 188-190. Could the authors elaborate a bit more on this processing step. What are “intervening cloud layers”? How is the continuity of a cloud layer evaluated? How is the cloud layer selected that the virga or precipitation is associated to?

19. P. 9 L. 191-192. Unclear sentence, I do not understand which part of the sentence refers to the cloud top and what to the cloud base values. Perhaps splitting the sentence to first describe the smoothing applied for the cloud top values, followed by a sentence describing what this is similar to, would help with to make more understandable.

20. P. 9 L. 193-194. I don’t understand the meaning of the sentence “This mapping is used to separate the cloud and virga-mask into cloud layer components”. What are “cloud layer components”?

21. P. 9 L. 195. Until here there has not been any differentiation between virga and precipitation, is the first step of assigning precipitation to virga to consider Z_e at the lowest radar range gate? What about multi-layer situations? Could the authors elaborate how from the cloud and precipitation mask (shown as a circle in the orange box 2 in Fig. 2) the virga mask is derived for the first time? And where in the processing (as described in Fig. 2) does this take place? Since in the orange box 3, and Sect. 3.3. (according to the subsection heading), the virga mask is refined, it appears as there should be a virga mask set prior to the (optional) third step.

22. P. 9 L. 195-198. Here the surface rain flag based on Z_e threshold is presented as part of the standard processing in step 2 (since it is described in Sect. 3.2), although Fig. 2 suggest it is part of the optional virga detection refinement in the orange box 3. Could the authors clarify the optionality of this processing step and where in the processing flow, as described by Fig. 2, it takes place?

23. Figure 3. This figure and the associated text are very nice and helpful for the reader to understand the details of the algorithm. Technically, the blue and pink values would also be valid Z_e values, the authors could consider using another label for the green boxes. The authors should also check that the figure is readable for colorblind readers.

24. Figure 3. Are the range gates intended to correspond to certain range resolution (so that i.e. allowed gaps would correspond to specific thresholds), or is the figure merely illustrative? An additional note on the caption would avoid ambiguity.

25. Figure 3. The figure clearly illustrates input and output parameters, however it is not clear to me to which part of the algorithm, as illustrated in Fig. 2, the processing illustrated by the Fig. 3

and associated text refers to? The entire orange box 2? Perhaps the authors could clarify which element of the processing chain the figure is illustrating.

26. P. 10 L. 200. No mentioning of a minimum virga length to be required has been provided until here, and it is also not included in Fig. 2. I found the explanation on the next page in the next section for virga mask refinement. Could the authors clarify where this criteria is used (step 2 or 3 of the processing), and on the optionality of this criteria?

27. Related to the previous comment, for the reader it would be less confusing if any criteria used for Fig. 3 and its description on lines 200-208, would be described prior to the figure and the text appearing. I therefore suggest the authors move Fig. 3 and the associated text later in the manuscript, when all parts of the algorithm used have been introduced, or move the explanation of the minimum virga length to before L. 200.

28. P. 10 L. 200. Could the authors either comment on rg 19 here, or remove this Z_e valid value in the corresponding time step in the figure?

29. P. 10 L. 206. In time-step 5 (and 6) there is no valid Z_e value in the lowest range gate, so obviously the radar reflectivity threshold could not cause the surface rain flag to be set. Perhaps the authors could add a time step, or edit time-step 5, to have valid Z_e-values to reach to lowest range gate, to illustrate the behavior in such a case?

30. P. 11 L. 213-214. Is the Z_e threshold at the lowest range gate an optional processing step, as suggested by Fig. 2, or is it always performed, as it appears from the text in Sect. 3.2?

31. P. 11 L. 218. How does the V_m threshold of 0 ms^{-1} perform in convective situations? Figure A1 suggests that for Z_e < 0 dBz, e.g. drizzle, V_m peaks very close to zero and values slightly above 0 ms^{-1} could be assumed to be drizzle observed in an updraft. Could the authors comment on the choice of this threshold in the context of convective situations, and have the authors evaluated the sensitivity of virga detection on this threshold?

32. Does the movement of the platform have an influence in the use of V_m -based virga refinement and the used thresholds?

33. Figure 4d. The contrast between the red and orange is quite poor, could the authors consider another choice of colors to aid the readability?

34. In Fig. 4a and 4c there is a line around ~300-400 that looks rather strange. Is this an artifact? Could the authors comment?

35. Figure 4 Caption. The caption is missing the mentioning of the radar reflectivity factor shown in panels a, c, and e, which I understood to be the input for the Virga-Sniffer.

36. Figure 4. What is the filled cloud base in Fig. 4b, d, and f, shown with a dashed line? Is it denoting the interpolated values (L. 171)? How come is the lowest cloud layer continuous, is this from the LCL filling (L. 163)? Clarification from the authors to correctly interpret the figure would be appreciated.

37. In the plots illustrating Virga-Sniffer results (Figs. 4, 5c), the green and blue are very hard to distinguish from each other, and I ask the authors to consider using colors with more contrast.

38. P. 15 L. 288. The authors mention here a smoothing at precipitation edges performed by the Virga-Sniffer algorithm. Could they please include a description of this procedure in the algorithm description?

Minor comments

39. P. 1 L. 16, 18. There seems to be a slight mismatch between the values presented in the abstract and those in Table 3, where the fraction of clouds below the trade inversion producing virga is 51% and the fraction of virga produced by trade wind cumuli is 37%, in comparison to 50% and 36%, respectively, written in the abstract. Could the authors correct this, or clarify where the values corresponding to those in abstract are to be found in the manuscript?

40. I believe the abstract as well as the conclusions should not make statements not supported by the results presented in the paper. The paper does not show the dependency of virga depth on liquid water path, only mentions that no dependency is found and the result is not shown. I recommend the authors either present the result, or remove the statement from the abstract. Furthermore, the paper does not provide any analysis on the cloud types producing virga, only relates virga to cloud height and depth. The attributions of certain virga features to cloud type are claims by the authors, not supported by any analysis in this study or references to the literature. Although these claims may be reasonable and supported by knowledge of the features of certain cloud types in the specific climate zone studied, I find it questionable to present these claims in the abstract and conclusions. I suggest that the sentence “The most important virga-producing clouds were either anvils of convective cells or stratocumulus clouds.” (P. 1 L. 17-18) be removed from the abstract, together with references to specific cloud type in the conclusions (P. 20 L. 372-373). Similarly, for the statement that virga detected by Virga-Sniffer that is classified by Cloudnet as ‘aerosols and insects’ occurs mostly at virga edges, the authors only show one case as evidence (Fig. 5). While I have no reason to doubt this result, the manuscript does not demonstrate that most cases are indeed like the one example shown, and I therefore suggest the authors consider if the statement should be included in the abstract (P. 1 L. 14-15) and in the conclusions (P. 19 L. 361-362).

41. I kindly ask the authors to add a note in the abstract that the results for virga occurrence reported are for the winter (dry) season.

42. P. 2 L. 32. “more numerous and smaller” → more numerous and smaller compared to?

43. The introduction well motivates the relevancy of studying the evaporation of precipitation in the trades. However, observation-based techniques used to detect or evaluate evaporation in the literature are not described. I kindly ask the authors to add some background on the observational techniques used in previous studies, given that the papers main contribution is to improve on the observational methods available to study precipitation evaporation.

44. P. 4 L. 99-101. While I agree with the authors that evaluating the performance of the radar stabilization platform is a topic for a separate manuscript, it would be relevant to comment whether there is an impact on the virga detection presented in Section 4.

45. P. 5 L. 104-105. Are there any references that could be provided for the radar data processing?

46. P. 5 L. 115 and 119. Integrated water vapor is not used in this study, and could be removed from the description of the data set.

47. P. 1 L. 15, P. 5 L. 129, and elsewhere in the manuscript. CloudnetPy is sometimes referred to as CloudnetPy and sometimes Cloudnetpy, consistent naming should be used.

48. P. 7. L. 136-137. How much data was removed due to radar settings not being compatible for CloudnetPy? In Section 2.1 and Table 2 two chirp programs are described, are the measurements corresponding to these settings included in the analysis? If yes, what is the data that is excluded?

49. P. 7 L. 138. Could the authors specify which model or reanalysis product from ECMWF was used?

50. P. 7 L. 152. Sect. A → Appendix A.

51. Some small editing is required for Table 1, specifically:
    a. The table includes parameters not used in this study (spectral power, spectrum width), which I suggest the authors remove. Alternatively, the caption should be edited not to specify that the table includes "measured quantities used in this study".
    b. For LIMHAT, the frequencies 22.23-31.4 GHz are missing.
    c. It is not clear what the two different temporal resolution, vertical range and vertical resolution values given for the LIMRAD94 refer to. Comparing with Table 2 they seem to be associated with the two main chirp tables used during the campaign. I ask the authors to clarify this, as the table as it is currently presented might lead to misunderstanding that Z_e and V_m have different temporal resolution and vertical range and resolution.
    d. For the ceilometer, I believe that the cloud base height should also be given as a Measured/received quantity, especially as I understand from the description that the instruments internal retrieval is used and cloud base height is the parameter given as input to the Virga-Sniffer.
    e. In the first row, I wonder if the authors would consider replacing *received* with *retrieved* in "Measured/received quantity", to reflect that some of the quantities (e.g. LWP) are *retrieved* from the measurements, and as I do not see the need to have both words 'measured' and 'received'.

52. Table 2 Caption. Parenthesis missing at the end.

53. P. 13 L. 254-255. Could the authors clarify whether the recommended virga mask refinement using mean Doppler velocity was used in Fig. 5? Also, do the authors here mean the clutter filtering as described by Eq. 1, or the 0 ms^{-1} threshold criteria, or both?

54. P. 13 L. 268. Unfortunately I don't see where radar reflectivity is connecting through multiple ceilometer observed cloud layers around 5:00 UTC, could the authors perhaps indicate this more clearly in the figure?

55. P. 13 L. 270. Could the authors comment on why they decided on ignoring the lower CBH when cloud layers are connected by precipitation, instead of the higher one?

56. Figure 6. It is difficult to gain quantitative information of the inner ring. Could the authors add some ticks (for example every 10%?) to give guidance, or label the largest blocks, to provide the reader better understanding of the results?

57. Figure 6. It is not obvious from the figure legend how the individual Cloudnet target classes are grouped into the liquid-only and ice-containing groups. The authors could add more information in the caption or the labeling in the figure. The authors could also consider

combining Figs. 6 and 7 to one figure with two panels, which would solve the problem, since Fig. 7 unambiguously shows which Cloudnet target classes are included in which grouping.

58. Table 3. The text states that only clouds with CBH below 4 km are considered in the analysis. Could the authors also mention this in the table caption, in case it is valid for the table, to avoid any misunderstanding to what the percentages presented refer to.

59. Section 4.2.2-4.2.3. Is virga depth computed here as the geometrical depth from the lowest to highest bin of the bin, or are gaps ignored? Would the different way of calculating the virga depth have an impact on the results and their interpretation?

60. Figure 8. The smallest virga depth bin seems pronounced, do the authors have an explanation for this?

61. P. 17 L. 324. Figure 8 → Figure 9.

62. P. 17 L. 329-330. Could the authors elaborate on which basis they are making statements about specific cloud types based on CBH, perhaps by adding some references to the literature?

63. The authors evaluate how cloud macrophysical properties, namely cloud base height, cloud depth and LWP relate to the virga depth. However, the sub-cloud relative humidity is quite relevant when considering the evaporation of rain. Could this provide some explanation why there are no strong relationships found in Sect. 4.2?

64. P. 19 L. 351-352. Similar to comment 40, I urge the authors to be careful to present results in the summary and conclusions section that were not actually shown in the paper. I suggest removing the statement about applying the Virga-Sniffer for RV *Maria S. Merian* measurements, because these are not shown.

65. P. 19 L. 358. Here 30% of virga detected by the Virga-Sniffer for the EUREC^{4}A data set is said to be classified by Cloudnet as ice-phase precipitation, but in Section 4.1. it is stated that 31% of virga pixels are classified as ice-containing (P. 15 L. 284). Could the authors explain the discrepancy?

66. P. 19 L. 363 – P. 20 L. 375. I don't see the purpose of the two one sentence paragraphs, and it seems to me that these two sentences could be merged to the following paragraph. However, I leave it to the discretion of the authors how they choose to present.

67. P. 20 L. 393. Are the default values given also the same as used for the processing the EUREC^{4}A data set in this paper?

68. P. 21 L. 424. Does this threshold correspond to the amount of valid data points required, as explained on P. 7 L. 154?

69. P. 22 L. 426. Which preprocessing? Does this refer to the CBH preprocessing described in Sect. 3.1? Could the authors give a more clear reference.

70. Figure A1. Colorbar is missing and should be added.