Review of "RainForest: A random forest algorithm for quantitative precipitation estimation over Swizerland"

By Irene Crisologo

Summary:

The paper presents a new method for Quantitative Precipitation Estimation, by using the Random Forest algorithm to convert radar variables into precipitation estimates. The authors created multiple products using the Random Forest algorithm, by using different inputs such as dual-pol variables, horizontal-polarization variables, and bias-corrected data. Their results suggest that Random forest show promising results for QPE.

Overall the paper is well written, the methods are presented well, and the results are conclusive. Just a few clarifications are needed, and some minor things edited for consistency.

General Comments

1. The inclusion of computational methods, a written algorithm, and citation of programming libraries are greatly appreciated.
2. The paper (understandably) uses several acronyms, and as they are defined in various locations in the text, it might help readability if an additional table defining acronyms is added. I understand that this is not a common thing to add to papers, so the authors may consider this optional.
3. The efficacy of the RF QPE is compared to a Z-R based QPE. Do dual-pol based QPE products exist that can also be used as a reference?
4. Table 1 is not referenced in the text.
5. What is the reason for selecting only 6 cases from 4 years of data?
6. The final products that the authors compared in the end could be presented in a more reader-friendly manner. The "RF_dualpol_AC" product for example is first mentioned in Page 24, and when I searched for the definition within the text, it is only defined in the caption of Figure 15 on Page 26.
7. Random forest is the main method and core idea of the paper, yet little about it has been mentioned in the introduction. Maybe a bit more information on Random Forest can be added in the introduction.
8. I would also like to suggest for the authors to check their figures against a colorblindness simulator for color schemes and choices that might not be very accessible.
9. I appreciate the vector figures, zooming in does not blur the image. However, the need to zoom in could be eliminated by having a consistent, readable font size throughout the manuscript. Some figures have font sizes that are too small. Ideally, the font sizes in the figure matches the font size of the manuscript body, or only a little smaller.

Specific comments

1. Page 1, line 13: "… that appears less smooth then the original…" change *then* to *than*
2. Page 3, line 26: COSMO is mentioned for the first time, but not defined. Please add a sentence briefly describing/introducing COSMO, as it is referred to in the later parts of the manuscript.
3. Page 5 line 10: Ott Pluvio footnote is missing.

4. Page 7, Table 2: The shorthand VIS could be added to the "Visibility" row of this table, since VIS and "Visibility" is used interchangeable throughout the manuscript.
5. Page 10, Figure 3: Maybe a monochrome colorscale could be used for the density plot? The green part of the colorscale slightly interferes with the green and blue lines in the plots. The colorbar is also missing.
6. Page 10, Figure 3: The legend entries are not defined anywhere. What do "raw_deg1", "cdf_deg1", and "cdf_spline" mean? Does the density plot in Panel (b) correspond to the green line in Panel (a)?
7. Page 12, line 14: What are the stations mentioned and where are they located? A bit more context about the mentioned stations could help.
8. Page 14, Figure 5: Much discussion is presented about Figure 5, it would help if Figure 5 is more readable and refined. The labels of the pie chart overlap, which makes it hard to read. NOISE_H and Zh has the same color of pie slice. Zv and RVEL also seem to be the same color. Some labels are also not defined anywhere, e.g. FRAC_RADAR. The variables also have inconsistent use throughout the manuscript, e.g. $Noise_H$ vs NOISE_H, RVEL vs R_{vel}
9. Page 16: There is a table but the table number and caption are missing. This table should also be referenced in the text.
10. Page 18, Figure 7: Is there a difference between "residual" in the legend and "daily" in the caption? Or do they refer to the same thing?
11. Page 19, Figure 8: Some labels overlap each other or the standard deviation indicator.
12. Page 23, Figure 13: What do the x- and y-axis represent?
13. Page 24, line 22: "RF_dualpol_AC" is used here before being defined in Figure 15 caption.
14. Page 26, Figure 15. The authors defined green colors as good performance, red colors as poor performance. However, there are multiple shades of greens and reds, as well as yellows. A legend or colorscale showing the thresholds used for the different colors and shades could be useful.
15. Page 26, Figure 15: The CPC.CV column seems to be the best performing of all methods, as it is mostly green. Discussion about this is missing in the text.

Technical corrections

1. Page 1, title: Swizerland -> Switzerland
2. Page 1, line 18: "Still providing" -> "Still, providing"
3. Page 6, line 2: Plane -> Plan
4. Page 10, Figure 3 caption: "Panel (a) shows ab example…" *ab* → *an*
5. Page 10, Figure 3 caption: Kindly clarify this part of the sentence: "…which are estimate the observations as a function…"

Very minor and somewhat nitpicky corrections

(I've recently read some books about data visualization and designing plots, and found this review to be a good opportunity practice identifying the little things that could improve a figure. However, they do tend to be nitpicky, so feel free to skip these suggestions!)

1. Figure 3
    a. The red 1:1 line could extend a bit more so that it goes from corner to corner of the plot, like it does in Figure 4.
    b. The colorbar is missing.
    c. The 1:1 line eliminates the need for the gridlines.
    d. The figure can be made bigger.
2. Figure 4
    a. The font size of the labels should be increased for readability.
    b. The aspect ratio could be set to 1:1 for each subplot
    c. The legend box alignment can be improved (my suggestion would be to place it on the top right, where it interferes least with the data)
    d. The figure can be made bigger.
3. Figure 5
    a. If possible, pie chart slices should appear in decreasing size, so that the order of slices (whether CW or CCW), has some logical sequence.
4. Figure 7
    a. The assignment of colors for the (a) Daily test RMSE and the Residual line in (b) Decomposed daily test RMSE makes it seem like they are the same variable.
5. Figure 9
    a. The red 1:1 line could extend a bit more so that it goes from corner to corner of the plot.
    b. The legend box alignment can be improved.
6. Figure 12
    a. The grid lines can be eliminated.