Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC1
https://doi.org/10.5194/amt-2021-175-RC1, 2021

Reviewer's comment on amt-2021-175

Anonymous Referee #2

Referee comment on "Leveraging machine learning for quantitative precipitation estimation from Fengyun-4 geostationary observations and ground meteorological measurements" by Xinyan Li et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2021-175-RC1, 2021

The paper "Leveraging machine learning for quantitative precipitation estimation from Fengyun-4 geostationary observations and ground meteorological measurements", by Li and co-workers, presents a preliminary application of a machine learning technique to retrieve precipitation hourly rate from geostationary VIS-IR data. A Random Forest classificator is applied to multispectral AGRI data on board the Chinese FY-4 satellite for 3 2-day storms occurred in Southern China: calibration and validation of the estimates are performed against hourly automatic weather station data.

The paper is interesting since there is very little published work on FY-4 data, however, I think the present manuscript needs a deep revision before to be published on AMT. Below, my suggestion to improve the quality of the manuscript.

Introduction

Lines 51 and following: any introduction on multisensor precipitation estimation cannot forget international programmes that provides high quality and high resolution precipitation products at global or continental scale, such as NASA-GPM or H-SAF. Please, complete.

Line 85, and in many other parts of the paper, are mentioned high-density stations, without any quantitative indication on how the density is measured and how "high density" is defined. Please, give more quantitative details on the station distribution.

Line 96. In Figure 1 please write the meaning of red areas (NPP_NTL?).

Lines 106 and line 108 mention levels: "met the levels for large-scale heavy precipitation" and "met the heavy rain level". How are these levels defined?

Data

Line 128, very likely, 4km is the nominal resolution at nadir.

Lines 145-154. ERA5 fields come with significant latency (5 days for the "preliminary daily updates" and 3 months for the "quality-assured data". Are these times compliant to the Authors' aim to "monitor flood" (line 81)? Moreover, how could be possible the "the real-time monitoring and prediction of summer precipitation over East Asia" (lines 383-384)? Please discuss the temporal applicability of the proposed technique.

Lines 155-171. Please, improve this description. First try to clearly separate different steps of the algorithm (e.g. with bullet points), then use different fonts to define variables in the text (mtray, ntray...).

Line 184 and elsewhere. Please, do not use the word "prediction" here and in the whole document to refer to the output of your algorithm, use "estimate", instead.

Lines 195-200. Here is the main lack of the paper: POD and FAR cannot be used separately to assess the quality of an estimates. Besides an error in the sentence ("optimal value of FAR is 1, and the worst value is 0", actually, the opposite is true), to measure the capability of the technique to correctly classify wet/dry pixel you need or to comment both POD and FAR number together (and avoid sentences as on line 17), or to compute synthetic indicators such as Equitable Threat Score (ETS), Hanssen and Kuiper or Heidke Skill Score, and do again the analysis looking at the values of these indicators as reference.

Results

Line 205, figure 4, please, use a reasonable number of digits in the numbers reported on the panels. Moreover, POD and FAR should be <1.

Lines 214-215. This sentence is a speculation not supported by evidence, please motivate it better or cancel.

Lines 218-221. To better illustrate this issue, please, use the indicators I suggested few lines above (ETS, HK…)

Line 223. Again, an unsupported sentence, please, give evidence or remove.

Lines 224-230. This paragraph not clear: if the Authors have the feeling that the dataset is not large enough to carry on proper training/testing procedure (that was also my feeling at the beginning) why not to add some more case?

Lines 245-246. I do not see this sentence comes out. Numerical indicators (POD, FAR, R, ETS….) tell us much more from the quantitative point of view with respect to simple visual comparisons of rain maps. Please, use numbers if you want to make quantitative assessments.

Lines 263-264. Absolute and relative errors are not defined in the text.

Lines 291-293. This sentence does not tell anything about the technique accuracy, since POD alone is considered.

In general, discussion and conclusion have to be rewritten once the new indictors I suggested will be implemented.