Interactive comment on “Assessment of the total precipitable water from a sun-photometer, microwave radiometer and radiosondes at a continental site in southeastern Europe” by Konstantinos Fragkos et al.

Anonymous Referee #2

Received and published: 29 October 2018

This study is very interesting, fitting the scope of ATM journal and could prove very useful for future studies in the field. Intercomparison of TPW retrieved from different techniques, for decade long timeseries is not found regularly in the literature and at this work data from radiosondes, sunphotometer and microwave radiometer are compared at such interval. Authors aim more to compare two versions of AERONET retrieving algorithm, which is very interesting but the results found are not discussed in depth, differences and uncertainty is not justified in depth. Thus, I suggest to consider the manuscript for publication after undertaking major revision. More specific com-
a) Various retrieving approaches have been studies in other publications, but authors should make clear which approach is considered more representative of actual atmospheric conditions. To my knowledge, radiosonde retrievals, being real in situ measurements at different heights, are the data that should have this role. Although there are uncertainty at this retrieval. In my opinion all comparisons and explanation should be performed according to the principle that more reliable are remote sensing retrievals closest to radiosondes. b) To my knowledge, there is no other publication for the version 3 algorithm of AERONET. Thus, a more detailed discussion on this algorithm is needed. Especially at paragraph 2.1, formulas and hypothesis used for retrieving TWP should be discussed. A very important aspect, is that algorithm uses AOD retrievals from other wavelengths, hence the differences to AOD between the two versions, should be propagated also to TWP. Also, some discussion is needed on the uncertainty of this retrieval and if there are any differences to it between the two versions. c) I have very serious doubt on the data from radiosondes and microwave radiometer. Sunphotometric retrievals usually are up to around 40 mm, because higher loads are usually linked to the presence of clouds. Hence, radiosondes and microwave radiometer that are measuring no matter the cloud conditions should have significant higher average and median values (table 1, figure 1). Either authors have used only synchronous to aeronet data for all retrievals, thus cloud conditions are filtered out or there is some severe problem with the datasets. This should be clarified in any of those cases and rewrite this paragraphs to clarify the procedure or recalculate all statistics. At p5 l4, it is stated that for daily mean values all measurements are used. d) More information on the climatology of the measuring site is needed at paragraph 2. I would suggest at least some statistics on yearly sunshine hours (which affects the quantity of aeronet data) and some range and averages for water vapor or at least moisture in the area. e) Since the normal distribution is visualized on histograms (red line - figures 5,10,12), it is expected to perform some statistical test to determine whether data’s distribution fits to it. I suspect from the plots that it doesn’t fit, so it is preferable to find the distri-
bution that best fits the data and at least some discussion in the manuscript should be expected in the manuscript. f) At paragraph 2.3 some discussion on the spatial spread of radiosondes should be added. I would suggest to filter out some radiosondes from the comparisons, with some criterion about the final position or at least the position at 4km height, which could make data uncomparable, since the distance between the two sites is already 30km and in case of southwest winds could be even larger for the sounding. g) Paragraph 3.2 needs clarification. What “randomly selected days” mean? How the random process was performed? What does “enough CIMEL observations” mean? Please be more clear when describing these procedures and make clear what conditions were applied and how was selected. h) P6 l10. This sentence should be more clear and have a more extensive discussion. What sza corresponds to these airmasses? Is that the airmasses that differs between v2 and v3? Higher sza values at 8 June and 26 October should be a lot different, are these differences observed at the same sza’s at these days? Is that pattern observed at other days at these angles? I would suggest to add a plot in figure 3 with sza at x-axis to make all this clearer. i) p6 l13. I would suggest to use the more robust approach found at Schneider et al. (2010) averaging measurements for ± 20 min from the time that the radiosonde reaches a 4 km height, in order to minimize spatial and temporal measurement differences. áÅÍ j) paragraph 3.4. I strongly suggest to investigate the differences in respect to AOD at 870 nm. Following the earlier about the missing methodology for TWP retrieval, I suspect that measurements with high differences in AOD between v2 and v3 will propagate to TWP values, especially at values less than 10mm, where the AOD influence is a lot larger. k) P8 l5, so it is suggested that v3 has slightly less accordance to the more reliable measurement. I suggest to use the Schneider criterion for averaging data around radiosonde, to have a more robust estimation and also I think that AOD differences will partially explain this behavior. Otherwise, it would be an interesting finding that v3 downgraded the quality of TWP. l) P8 l32. This trend is calculated statistically, but for lower values there is a very high spread of differences, thus I have doubt if this statistics is meaningful. Practically values below 10mm could have any difference, and
higher values converge a lot. This behavior is explained through the uncertainty of both instruments that lowers for higher TWP. This should be discussed in respect to uncertainty estimations. m) P19 l15. What is the distance to radiosonde launching site? 18 or 30 km (stated in paragraph 2.3)
