I struggle with the assessment of this work: while the all the methods and results presented are of respectable scientific quality, I think there is a lack of focus in terms of relevance and scope for AMT. There is no novelty in concepts or data treatment and it is not clear what the real scientific value of the study is. For regulatory purposes there is little value due to the non-compliant sampling system, non-existent pre experiment calibration etc. The scientific value is also limited – I understand the argument for connecting ground measurements to cruise at altitude data, but for that purpose, a more focused effort with a better experimental design that would allow tracking down sampling/conditioning from instrument issues would be beneficial. With the current manuscript one gets the impression that it is a side product of a bigger effort and was not carefully thought through when the experiment was conducted – which is not necessarily a problem if the reader does not get this impression, but I currently do.

Major comments:

- The comparison of the mass measurement is somewhat biased experimentally (due to distance to the engine, dilution, detection limits and long lines etc.) to higher thrust levels. At these thrust levels it is not a major surprise that there is not much variability in instrument responses (little OC, larger aggregate sizes, soot properties less influenced by fuel type etc.). I also tend to disagree with the authors conclusion that a 30-50% difference is a “comparable” especially for the near real time in situ instruments such as MSS LII and CAPS. Would be good to point this out to the reader, or even split the discussion for cruising relevant (i.e. 50 -70% thrust) and near idle thrusts this might improve the lack of relevance pointed out above.
- It would beneficial to show the comparison of measured concentration as a function of CO2 (at least in the SI)
SMPS EIm derivation: this work makes the impression that an SMPS measures the volume size distribution with high precision and there is furthermore no need to apply a size dependent effective density (which I believe is crucial for larger sizes). It would be beneficial for the discussion to elaborate on this based on previous experiences on helicopter or jet engines.