Comment on amt-2021-360
Anonymous Referee #1

Referee comment on "Air quality observations onboard commercial and targeted Zeppelin flights in Germany - a platform for high-resolution trace-gas and aerosol measurements within the planetary boundary layer" by Ralf Tillmann et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2021-360-RC1, 2021

Tillmann et al. present observations of greenhouse gases and important primary and secondary pollutants that impact air quality. These measurements were conducted on a Zeppelin, allowing for both targeted measurements as well as measurements conducted at lower altitudes than typical for aircraft. With the observations, the authors were able to show the collapse of the nocturnal residual layer and the mixing of the boundary layer with the residual layer, some investigation into emission sources, and comparisons of emissions/transport with a model for selected sources.

Though the paper is of interest and provides valuable results/information, currently as presented, the paper seems to be more geared for Atmospheric Chemistry and Physics and not Atmospheric Measurement Technique. The reasoning behind that is discussed in more detail below. Further, there are areas where further clarification and discussion about the methodology is needed. Depending on what direction the authors decide to take will make this paper either acceptable for AMT or ACP.

As this was submitted to a measurement technique journal, it would be expected that more details would be included concerning the measurement and techniques. However, the authors provide minimal information about the measurement techniques. The following, at minimum, should be included to make this paper more appropriate for AMT:

1) Comparison of the low-cost sensors along with MIRO. Including, but not limited to, sensitivity, response times, interferences (including any potential RH or temperature dependencies, esp. if this unit is not in a temperature controlled area of the Zeppelin), any pressure dependencies, how limit of detection may change with integration time, etc.
2) Response time of MIRO, along with the lines used. This is especially important as:
   a) MIRO is measuring sticky molecules (H2O and NH3) so how quickly can MIRO actually
      respond to entering/exiting boundary layer?
   b) As one of the selling points of this package is that it can better characterize boundary
      layer, residual layer, and potentially free troposphere, the instrument quickly responding
      entering & exiting these different regimes is important. This can also have potential
      relative humidity dependencies that will be important to discuss and outline.

3) Discussion of the sampling scheme--what is the tubing (type, internal diameter, length)
   from the inlet to the different instruments? What is the residence time? Is it heated?
   Similarly for the zero air supply?

4) What is the cooling system for? What is the zero air catridges for? What is the pressure
   controller for? What happens without these devices?

5) How reliable is this instrument on maintaining the lines/features to measure the
   compounds listed? How easy is it to get back to the lines/features if the instrument loses
   them due to pressure or temperature fluctuations?

6) An important question that the authors brought up includes boundary layer and
   residual layer. Currently there are minimal measurements in the residual layer and in
   differentiating boundary layer, residual layer, and free troposphere. Is there an alogarithm
   from the observations the authors presented that they can estimate the heights of these
   levels?

The following clarifications/discussions would improve the paper:

1) It is currently unclear throughout Section 3 if the authors are only talking about MIRO
   or if it is a combination of MIRO and electrochemical sensors.

2) As a side note, since the authors brought up the electrochemical and optical
   measurements, at least a minor discussion of these measurements would be beneficial for
   this manuscript. Further, for the optical counter, what type of line (Teflon vs
   copper/stainless steel) and dryer used?

2) As geostationary satellites are coming on-line, a discussion in how this package could
   be used to validate the geostationary satellites would be of use. This is especically
   important as the package includes many of the species that the satellites will be trying to
   target. How does the integrated column change between the different times the authors
   discuss, along with the air mass factor, between having a nocturnal residual layer and a
well mixed boundary layer?

3) The purpose of Fig. 4 and the associated discussion is currently not clear. Numerous studies from NASA DISCOVER-AQ and KORUS-AQ have indicated that NOx is not well mixed between ground and at altitude measurements (e.g., Flynn et al., 2014; Flynn et al., 2016; Choi et al., 2020; Li et al., 2021). A discussion in how the profile from the ground-based observations and the Zeppelin profile may be more suitable.

4) Why does it take the NOx concentrations to be higher at 250 - 300 m longer than rest of the boundary layer? Does it make sense with mixing/collapse of the residual layer?

5) The authors mention that compounds need to have similar loss rates in order to do ratios to understand emission ratios/sources (Section 3.3). However, NO2 has a very short lifetime (shorter than the 7 - 11 hours the authors noted due to the production of peroxy acyl nitrates, PAN, and alkyl and multifunctional nitrates, e.g., Valin et al., 2013). The authors should show what the ratio of NOx to CO from ground monitoring sites are and compare with the Zeppelin measurements to (a) show confidence in the assumptions and ratios they provide and (b) what new information the Zeppelin provides that ground measurements currently may not provide.

Minor:

1) Please make sure to be consistent about underscoring the x in NOx.

2) Please recheck the grammar and capitalization throughout the manuscript. E.g., World Health Organization is not capitalized and there are many instances were commas would be appropriate to separate a describer (e.g., line 50, "networks e.g. the European Environment Agency EEA together" should be "networks, e.g., the European Environment Agency, EEA, together"). Also, contractions (e.g., don't) should not be used.

3) It is generally recommended that references that use websites should have similar inline references as papers in that the website reference is listed at the end in the references section.

4) Why is the evening rush hour for NOx at such a late time (~8:00-9:00 PM local time). I would expect the rush hour to be between 4:00 and 7:00 PM local time.

5) Fig. 4: why are some of the flight paths thicker?
6) Fig. 7: I would recommend the observations have a black outline as they are hard to see with the background model results.