

Comments for the AMTD Manuscript, 'The Airborne Romanian Measurements of Aerosols and Trace gases (AROMAT) campaigns'

This paper discusses the air quality measurement campaign, AROMAT, and puts it into the context of if/how these measurements can assist in future validation efforts for satellites, such as Sentinel 5P TROPOMI. The significance of this work is within the scope of AMT and is key as the air quality community works toward validating satellites that measure urban air quality (e.g., TROPOMI) and there are some novel 'take-home' messages from this work that are worthy for publishing. From what is shared in the paper, the quality of the work appears valid however there is clarity needed in some areas. This paper also needs restructuring to improve the clarity of the take-home conclusions. For example: The paper lacks details about the campaign and information about the measurements are scattered throughout the paper.

Specific comments/questions:

- The title of this paper does not clearly reflect the contents of the paper. The current title would attract readers as an overview paper for the measurements during AROMAT, but this is not the purpose of this paper nor is there a detailed overview of the entire campaigns. If agreed by the authors, I suggest changing to title to something that reflects that AROMAT could be a concept model for validation campaigns of satellite retrievals.
- Are conclusions made about validation only valid over Romania? Or can these lessons be extended beyond Romania? Please be clear in the paper which conclusions can be extended beyond the AROMAT region.
- It seems that the model for the conclusions is based on TROPOMI requirements. Please comment on if/how this extends to the requirements of the other planned missions or make the specific message in the paper that the conclusions that are made are specific to TROPOMI.
- A weakness in the general analysis is the lack of discussion on temporal variation and the time of the airborne and ground based measurements and how this relates to the time of the satellite overpass, emissions inventories, etc. The authors should keep this in mind to address through revisions.
- Section 2 should start by painting a picture of AROMAT 1 and 2 deployments and measurements that used in this analysis. While much of this information is in the supplement and scattered throughout the paper, the general reader enters Section 3 without the proper background to assess what is being discussed. Currently, it does not effectively communicate the needed details about the AROMAT campaign before moving into the results sections. To fix this, the authors could reorganize the section by moving 2.3 to before 2.1 and 2.2. Then there needs to be discussion (and maybe a Table) that summarizes each campaign. This table and/or discussion must include time periods of each deployment, location of each deployment, payloads for the aircraft and relevant details about ground measurements (in line with Tables 4-6) and could extend into partners and other details from 2.3 as seen fit.
 - Table 3 does not add substantial information to this paper and that space would be more effectively be used to summarize the campaigns themselves.

- Section 3 and 4 are hard to follow as it jumps between regions and different trace gas measurements. A suggestion would be to reorganize into sections focuses on specific trace gases. For example: Section 3 could just be about NO₂. With the following sections.
 - 3.1: similar for 3.1.3 with summarizing Bucharest observations
 - 3.2: similar for 3.2.1 with summarizing The Jiu Valley observations
 - 3.3: Relevant discussion from Section 4 about lessons about validation
- Section 3.1.1 and 3.1.2 and their associate figures do not fit within the scope of the paper as separate sections. Any relevant discussion could fit in within the other trace gas sections, in the supplement, or omitted.

The below of the comments are organized by trace gas.

- NO₂:
 - Line 209: The statement about the datasets in Figure 7 appearing consistent is not valid, which is alluded to later in the paragraph. Please reword or omit that statement in the discussion.
 - Line 221: what is the difference in time between the two measurements?
 - The statement about NO₂ vmr at 300m being a proxy for NO₂ VCD is not valid. It may be for that specific case but not overall. The results over the Jiu Valley even refute this statement.
 - Line 229: Is the Avantes spectrometer the Bremen Nadir instrument from Table 7? Please make descriptions consistent.
 - Not required but Figure S3 seems like a good candidate to move to the actual manuscript for comparing/contrasting with SO₂. It could also be helpful to see how Figure 7 and other airborne figures translate to the TROPOMI pixels. When talking about the validation context.
 - Line 334: It seems that temporal variation could also lead to overestimation in the slope depending on how the NO₂ is varying through time.
 - It should be noted that the temporal variation uncertainty quantified in this paper was specific to that area during that particular morning and more data would have to be analyzed to see if this is a typical value or not. These temporal variations are also likely much different during the time TROPOMI overpasses (not in the early morning) and on different days. Though the technique for quantifying temporal variation using the airborne data is novel and would be interesting to extend to other datasets.
 - It would also be helpful to add some more details in the writing or references about the exercise done in the first paragraph of section 4.1.3 so it can be recreated by others with similar datasets.
- H₂CO and SO₂
 - Line 239-240: Are there H₂CO direct emissions in Bucharest? That seems to be the implication with the statement in line 239.

- This is relevant to both SO₂ and H₂CO since they both have the conclusion that validation of satellite H₂CO and SO₂ is better suited with ground-based measurements.
 - Is this a recommendation only for Romania?
 - Are there ground based measurements from AROMAT that can be discussed in terms of validation like the airborne data is? If so, add this to the discussion. If not, the conclusion that ground-based measurements would be better suited than airborne may not be valid.
 - Its mentioned that the individual flights cannot always help in validation, which is true. But systematic measurements may help as discussed in the conclusions. Please say something about this within the sections themselves.
- Emissions section (section 4.4):
 - This section lacks sufficient background on the methodology for computing fluxes and lacks the context on how this fits with the scope of the paper. Fluxes are not mentioned in the abstract, intro, nor is emission estimate validation within the requirements for validation of satellite products. Though emission estimations put into the context of satellite applications is important and evaluating that is very important scientifically from that perspective.
 - Options:
 - Omit this section.
 - Add sufficient details or references for emission flux calculations and put this into the context on how this helps with satellite product evaluation as alluded to in the latter part of section 4.3. (Though, this section with all details could potentially be a stand-alone manuscript).
 - If kept, when comparing the emissions to inventories, be sure to consider variations in emissions from hourly/daily/seasonal timescales and the AROMAT measurements were only a small subset in time.

Other comments within the text:

- Line 27-28: Veefkind et al., 2012 doesn't reference the 3.5x5.5km resolution. Refer to the switch through the Readme file or another reference that talks about it: <http://www.tropomi.eu/sites/default/files/files/publicSentinel-5P-Nitrogen-Dioxide-Level-2-Product-Readme-File.pdf>
- Line 45: Can it be made clear what small signals mean? Does this mean the small signal:noise ratios or more a reference to clean areas that don't have a lot of signal?
- Line 114: what are the European thresholds? Add a reference and values.
- Table 1: change GEMS to launched instead of planned.
- Throughout the paper: Spatial resolution is in km and not km². For example, 7 x 7km² is not the same as 7km x 7km.
- Line 261: delete the word 'those'
- Line 396: change 'As for' to 'Similar to'
- Line 399: Start a new paragraph with the sentence starting with 'On the other hand'