Reply to the comments from referee # 2

(In red our replies, in italic the text already written, in green the added text.)

We first kindly thank the referee for his time, useful comments, and constructive criticism. We used his suggestions to prepare a new version of the manuscript.

The manuscript titled The Air borne Romanian Measurements of Aerosols and Trace gases(AROMAT) campaigns at two areas provides relevance of each instrument for validation of air quality satellite (e.g., TROPOMI) products. The paper identifies a significant source of comparison error (measurement time difference), which is a useful information for the satellite validation. It summaries DL, BIAS, measurement range of several trace gas species for each instrument. However, the paper misses detailed description of instrument characteristics and measurement geometry, data used for each instrument AMF and their effects of the retrieved products. There has been no analysis about horizontal and vertical representativeness of each instrument although the campaign is to aim for validation of TROPOMI. The manuscript needs to be improved considering those major issues.

There were already two published studies (Meier et al., 2017, Merlaud et al., 2018) dedicated to the AROMAT airborne measurements, these studies include the AMF description and vertical sensitivities (box-AMF) of the airborne DOAS instruments. The Supplement already gives technical description of each instrument, giving the published references.

We agree that these two papers and the Supplement were not visible enough in the manuscript so we added several references to it (see below). We also rephrased the end of the introduction:

Two aforementioned publications focused on the AirMAP and SWING operations during the 2014 AROMAT campaign (Meier et al., 2017; Merlaud et al., 2018). In this work, we present the overall instrumental deployment during the two campaigns and analyze the relevance of these measurements for the validation of several air quality satellite products: tropospheric NO_2, SO_2 and H_2CO VCDs.
[…]
The paper is structured as follows: Section 2 describes the two target areas and the deployment strategy. Section 3 characterizes the investigated trace gases fields in the sampled areas. Section 4 presents a critical analysis of the strengths and limitations of the campaign results while elaborating on recommendations for future validation campaigns in Romania. Eventually, we use the AROMAT measurements to derive NO_x and SO_2 fluxes from the two sites. The Supplement presents technical details on the instruments operated during the campaigns and presents additional information and measurements.

We also added a schematic for the geometry of the measurements.

About the horizontal representativeness, we emphasized in the conclusions that one main advantage of continuous airborne mapping is that the horizontal representativeness error cancels.

Abstract and Introduction: The objectives of this present study and campaign needs to be clearly distinguished. The objectives of the campaign are described in Abstract as "Their main objectives were to test recently developed air borne observation systems dedicated to air quality studies and to verify the concept of such campaigns in support of the validation of space borne atmospheric missions such as the TROPOspheric Monitoring Instrument (TROPOMI)/Sentinel-5 Precursor (S5P)." However, there are differences between the objectives of the campaign and those of this present work. Please address the objectives of this present study in Abstract.

We agree with the comment and we have added the objectives of this paper in the abstract

We present the AROMAT campaigns, focusing on the findings related to the validation of tropospheric NO2, SO2, and H2CO. We also quantify the emissions of NOx and SO2 at the two sites.

The objectives were already described in the introduction, but we rephrased it to better define the scope of the study (see above).

Line 218: ".The comparison reveals a good agreement when averaging the forward and backward-looking Mobile-DOAS NO2 VCDs, with a MPIC/AirMAP slope of 0.93 and a correlation coefficient of 0.94." One of the campaign objectives is to identify relevance and capability of each measurement type on ground or air borne platforms for validation of TROPOMI products. There are missing of both qualitative and quantitative causes for "slope (between ground based MPIC mobile DOAS and AirMAP) of 0.93 and a correlation coefficient of 0.94".

We are comparing collocated and almost time coincident measurements of NO2 VCDs. If everything would be perfect, the slope and correlation coefficient would both be 1. The remaining difference is small and may have several causes: instrumental bias, the small time difference of the measurements, errors of AMFs, different horizontal sensitivity. We have added this in the text as:

The remaining discrepancy may be explained by AMFs errors and differences in time and horizontal sensitivity.

Line 224: I do not understand how "the NO2 vmr measured at 300 m a.s.l. can be used as a proxy for the NO2 VCD". Please describe how it can be used used as a proxy for the NO2 VCD. Please also use capital letter for VMR rather than vmr.

We had tried to give the explanation in the next sentences, which compared the NO2 VCD derived from the proxy with the AirMAP NO2 VCD measurement but we agree it was not clear enough. We have rephrased for clarity and changed vmr to VMR here and across the document

This suggests that along this portion of the flight, which was inside the plume but outside the city, the NO2 VMR measured at 300 m a.s.l. may be used as a proxy for the NO2 VCD. Indeed, the BLH was about 1500m (Fig.S9 in the Supplement) during these observations. Assuming a constant NO2 250 VMR of 3.5 ppb in the boundary layer leads to a NO2 VCD of 1.4 x 1016 molec cm2. This estimate is close to the

AirMAP NO2 VCD observed in the plume (Fig. 6). When measured at 300 m a.s.l., the NO2 VMR thus seems a good estimate of its average within the boundary layer. Note that this finding is specific to the configuration in Bucharest where we flew at 10 km from the city center and does not apply to our measurements in the exhaust plume of the Turceni power plant (Fig. 9). Future campaigns should include vertical soundings inside the Bucharest plume to further investigate its NO2 vertical distribution.

Line255-260: In comparisons between data of airborne AIRMAP, SWING, and ground based Mobile DOAS, it is important to explain if they measure the same target in terms of horizontal and vertical coverage. -If each instrument measures a target (in particular plume) at different geometry and location, there should be large differences between the retrieved NO2 VCDs. Authors need to explain reasons that cause such differences in terms of the algorithms, measurement geometries, effect of platforms, etc., in detail.

The instruments aim at the same target at the same time but from different locations and geometry, the ground for the Mobile-DOAS (zenith-looking) and from 3 km altitude for the airborne-DOAS (nadir-looking).

We had explained the main reasons for these differences, to our understanding, at the end of the paragraph, which also refers to our previous study which already compared airborne and mobile-DOAS measurements:

This is partly related to air mass factor uncertainties, but probably also to 3-D effects as the plume is very thin and heterogeneous close the power plants, as discussed in Merlaud et al. (2018).

We have to invoke another reason (3-D effects) in addition to the AMF only, since the latter can not explain the discrepancy between Mobile and airborne measurements. At the time of writing our manuscript this was still a conjectural but colleagues from another team are studying that with a 3D RT code, and their results seem consistent with what we wrote. See e.g. this presentation at ATMOS 2018

Implementation Of Three-Dimensional Box-Air-Mass-Factors In The LibRadtran Radiative Transfer Model

Schwaerzel, Marc; Emde, Claudia; Kuhlmann, Gerrit; Brunner, Dominik; Buchmann, Brigitte; Berne, Alexis

http://atmos2018.esa.int/page_session11.php

We rephrased and added a sentence to strengthen our initial explanation.

This is partly related to air mass factor uncertainties, but they can not explain alone such a discrepancy. Close to the power plant, the plume is very thin and heterogeneous which leads to 3-D effects in the radiative transfer, as suggested in Merlaud et al. (2018). In these conditions, the 1-D atmosphere of the radiative transfer models used to calculate the airborne AMFs may not be realistic enough and bias the VCDs measured from the aircraft.

-In the paper, a difference between mobile DOAS and those of airborne is partly related to air mass uncertainties. There is absence of description of NO2 AMFs for mobile DOAS and those for AirMAP and SWING. What are the input data used to calculate each AMF?

We agree it was not clear enough. For the airborne instruments, the NO2 profile is a box of 500 m as used in the reference we give in Meier et al. 2017 for Bucharest during AROMAT-1. We agree it was not clear enough that it was the same so we added the Phd of Andreas Meier as a reference in the AirMAP section of the supplement. This PhD includes AirMAP operation during AROMAT-2.

Note that a PhD thesis Meier (2018) describes in detail the AirMAP operations and the algorithms used to analyze the AROMAT data.

For the Mobile-DOAS, it was a zenith-only measurements and- we simply used the geometric approximation i.e. 1, as mentioned as a typical AMF value in the AMF NO2 table of the supplement. Here the AMF does not correspond to the reference (Constantin et al., 2013) since this previous work used a Chimere profile, which is not representative of the plume so close to the power plant.

We mentioned in the text and the legend of the figure that the Mobile-DOAS were zenith-only

Both AMFs actually correspond to the typical values given in the NO2 AMF table of the supplement, which we further emphasized:

Table S2 in the Supplement gives the typical AMFs used for this analysis for airborne and zenith-only Mobile-DOAS.

-please add schematic graph which shows instrument setup and measurement geometry (including measurement azimuth angles for target locations such as location of plume) of each instrument

We added a schematic (Fig.4) to explain the main campaign set-up. We did not add the azimuth however since as most of the measurements were mobile, it varied between 0 and 360°.

We have added a sentence presenting the figure in Sect 2.2

Figure 2 illustrates the typical instrumental deployment during the campaigns, which combined airborne and ground-based measurements.

This is the legend of this new figure 2 :

Geometry of the main measurements performed during the AROMAT campaigns. The Imaging-DOAS instruments map the NO2 and SO2 VCDs at 3 km altitude above the target area while the in-situ samplers measure profiles of trace gases and aerosols. Ancillary ground measurements include Mobile-DOAS to quantify trace gases VCDs and lidars to measure the aerosol optical properties.

Line 300: There are many sentences which mention “reference measurements”. Please define “reference measurements”

Following Richter et al. (2013), we mean “independent data with known and documented uncertainties” that we can meaningfully compare with satellite products. We have expanded the sentence in the introduction which first use this expression:

Validation involves a statistical analysis of the differences between measurements to be validated and reference measurements, which are independent data with known uncertainties (Von Clarmann, 2006, Richter 2013).

Line 304: What are “typical air mass factors (AMF) used here for each species and what are the references for each AMF value for each species for each instrument?

The sentence just after (line 305) already indicates the typical AMF: *“Table S1 in the Supplement presents these typical AMFs and detection limits”*. The references for each AMF comes from their different reference paper. We have added that in the legend of the three AMF tables.

See the references in Sect. 2 for details on the AMF calculations of the airborne instruments. We used geometric approximations for the ground-based DOAS instruments, pointing to zenith (AMF = 1), and 22° above the hrozon (AMF = 2.7).

Line 394: Please address the definition of “combined uncertainty” including how “combined uncertainty” has been calculated.

We had already explained this definition at the beginning of the sentence*: Adding in quadrature the biases of the SO2 VCDs for airborne measurements (40%, Table 6) and for TROPOMI (30%, Table 2) already leads to a combined uncertainty of …’ It seems already clear to us.*

Throughout the figures tables, there no quantitative comparisons between various measurement data which were carried out at the same or similar time in the same site. Please consider adding the plots with analysis or address the reasons for not doing that.

We had put the quantitative comparisons in the Supplement, where there are SWING vs AirMAP comparisons for NO2 and SO2 (Fig. S1 and S2) and AirMAP vs Mobile-DOAS measurements (S6), with correlation coefficients, slopes, intercepts, and number of points. We also show a CAPS versus NO2 sonde intercomparison (Fig. S5) but here it did not make sense to quantify the slope as the sonde was calibrated with the CAPS data. So we think we have already the quantitative intercomparisons needed to support the conclusions of our study. But we agree it was not visible enough in the main article so we added references to these figures in the section 3.2.2 and 4.1.2 where these corresponding measurements are discussed and used:

In Sect. 3.2.2, we had already written that the SWING and AirMAP VCDs agree within 10%, we have added:

Figure S4 in the Supplement shows the corresponding time series of SWING and AirMAP SO_2 DSCDs.

In Sect. 4.1.2

Figure S3 in the Supplement presents the corresponding AirMAP and SWING NO_2 DSCDs.

Moroever, we added a quantitative description in the conclusion for the comparisons between airborne and mobile-DOAS in Bucharest.

These measurements agree within 7\% *with ground-based measurements*