Response to anonymous referee #1

The reviewer comments are given in black, followed by the authors' response in blue. Text copied from the revised manuscript is in blue italic.

Based on a comment from referee #1, Fig. 8 and 9 have been merged and, as a result, the numbering of the following figures has been changed. In addition, the structure of Sect. 2 has been modified in order to include a new subsection with the description and implementation of the GRASP algorithm (Sect. 2.6).

General comments:

Overall, this is a well-written paper evaluating an interesting aspect like the effect of the columnar NO_2 correction in the accuracy of the Aerosol Optical Depth (AOD), the Angström Exponent (AE) or the Single Scattering Albedo (SSA) using multiannual records (5 years). This paper addresses some important aspects for the scientific community, such as the investigation of the effects of using different NO_2 data (no correction, direct retrievals or climatological values), which can impact the aerosol products retrieved with different global aerosol networks (NASA-AERONET, GAW-PFR or SKYNET-Prede). Moreover, NO_2 satellite data (TROPOMI) and spectral ground-based data (from Pandonia Global Network, PNG) were used to investigate the possible improvement in aerosol properties retrieved from these three largest ground-based aerosol networks. Trend analysis has been included to understand the impact of the NO_2 correction on the derived aerosol products, although the authors make it clear throughout the article that the number of data in the database is insufficient to carry out this type of study.

I consider that this manuscript fits perfectly into the scope of AMT and that the results presented here are relevant. There are only a few minor remarks.

We would like to acknowledge the referee for their helpful and thorough review. We believe that their comments improved the quality of this work.

Minor comments:

Abstract: Is uncertainty estimation a goal of this paper? I consider that this work deals with the impact of the columnar NO_2 effect rather than evaluating/investigating the uncertainty introduced by this term. Please state.

In principle, AOD uncertainty estimation due to NO_2 column used for AOD post processing is the main goal of the paper. The way to treat this theoretically would need an estimation on the used NO_2. Therefore, we have tried to demonstrate errors related with the use of climatological NO_2 rather than the actual measurements of NO_2. So a first step towards this goal is the evaluation of the satellite climatology used, based on collocated NO_2 observations from a state-of-the-art network. The second step is the assessment of these climatology vs measurements difference on AOD retrieval. For instruments not correcting for NO_2, it is straightforward that the AOD retrieval bias is directly linked with the NO_2 amount as in reality what is defined as AOD is the sum of the aerosol and NO_2 optical depth. Finally, as this study was performed on a pilot-test urban area, and can be, thus, considered “local”, we tried to assess the use of real-time satellite data for such corrections too.

So in our case the AOD uncertainty is always linked with the NO_2 related uncertainty, but here we present an idea of two-instrument synergy towards less uncertain NO_2 data, leading to less uncertain AOD data.

Page 2, lines 34-37: The authors are introducing the direct and indirect effects of aerosols. Don't the authors believe that there are more adequate references to introduce these effects? At least one more recent version of the IPCC exists than the one included in this article.

The authors thank the referee for pointing this out. The reference has been updated with a more recent version of the IPCC assessment report:

IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, In press, doi:10.1017/9781009157896.

Page 4, line 101: Is “specifically” right in this sentence?

This sentence is used as a more specific description of the two stations mentioned in the previous sentence. Thus, we think that the word “specifically” could be considered okay in this sentence. A comma separator has been added after this word so that the meaning is clearer:

“In this study, we used remote sensing measurements of columnar NO_2 and aerosol properties performed in two stations located in the greater area of Rome. More specifically, observations were obtained from an urban station…”

Page 4, line 112: The authors write “Cimel” in this sentence but “CIMEL” later in the text. Please homogenize.

The manuscript has been revised accordingly.

Page 4, line 118: Please note that Version 3 Level 1.5 includes data with near-real-time automatic cloud screening and automatic instrument anomaly quality controls while Level 2.0 additionally applies prefield and post-field calibrations. This means that the 1.5 level does not in any way apply the final calibration, so the lack of certainty in the verb “may” does not seem correct.

The word “may” has been removed from the sentence.

Page 4, Fig. 1: The authors present here a time evolution of the AOD and AE observations at APL-SAP and CNR-ISAC. I don’t see the point of including such a figure, because these data are not exploited here nor are they mentioned throughout the text.

This figure has been added to provide the reader with an overview of the aerosol data used in this work, and, in particular, to show typical values and seasonal cycles of the aerosol metrics addressed. We think it is a nice introduction to the AOD and AE levels in the two Rome sites. It is also useful as comparison to Fig.2, showing relevant NO_2 data, and, thus, highlighting the different seasonal cycles of the variables. For these reasons, we prefer to keep it. Following the suggestion, we included a brief description of the figure in Sections 2.2.1 and 2.2.2 as follows:

“The aerosol data sets for both locations are presented in Fig. 1. The average AE is 1.23 ± 0.4 and 1.31 ± 0.5 at APL-SAP and CNR-ISAC, respectively, while the average AOD is about 0.18 ± 0.1 at both stations. AOD has a quite marked yearly cycle, with higher AOD values recorded during summer months, i.e., about 0.22 ± 0.1 and 0.21 ± 0.1 at APL-SAP and CNR-ISAC, respectively. AE is also higher during summer with a mean value of 1.26 ± 0.4 for APL-SAP and 1.38 ± 0.5 for CNR-ISAC…

…The SKYNET time series used in our analysis is also illustrated in Fig. 1. The calculated mean AOD and AE are 0.18 ± 0.1 and 1.23 ± 0.4, respectively. These values are similar to AERONET APL-SAP averages mentioned in Sect. 2.2.1, though they correspond to slightly different wavelengths. SKYNET also reports higher values on average during summer, i.e., 0.22 ± 0.1 and 1.38 ± 0.5 for AOD and AE, respectively.”

Page 4, last paragraph: The authors have used level 1.5 SSA information from AERONET. However, as stated by Sinyuk et al. (2020), quality-controlled SSA data (level 2.0) should be retrieved for AOD larger than 0.4 and SZA larger than 50°. How the authors have ensured the quality of the SSA information included in this paper? Why the authors have not included AERONET Level 2.0 data in this study?

The levels of uncertainty provided by Sinyuk et al. (2020) clearly state the difference in quality of the different AERONET data levels in the mentioned conditions. However, these restricted conditions imply an extremely reduced amount of available data that makes impossible comparisons with a proper level of statistical significance. Thus, the authors consider that the trade between the amount of data and the loss of accuracy in the retrieved values results beneficially for the final quality of the comparisons. Also note that this methodology has been successfully applied in several publications as for example in Román et al. (2017), Román et al. (2018), Benavent-Oltra et al. (2019), Herreras et al. (2019).

Román, R., Torres, B., Fuertes, D., Cachorro, V. E., Dubovik, O., Toledano, C., ... & Alados-Arboledas, L. (2017). Remote sensing of lunar aureole with a sky camera: Adding information in the nocturnal retrieval of aerosol properties with GRASP code. Remote Sensing of Environment, 196, 238-252.

Román, R., Benavent-Oltra, J. A., Casquero-Vera, J. A., Lopatin, A., Cazorla, A., Lyamani, H., ... & Alados-Arboledas, L. (2018). Retrieval of aerosol profiles combining sunphotometer and ceilometer measurements in GRASP code. Atmospheric Research, 204, 161-177.

Benavent-Oltra, J. A., Román, R., Casquero-Vera, J. A., Pérez-Ramírez, D., Lyamani, H., Ortiz-Amezcua, P., ... & Alados-Arboledas, L. (2019). Different strategies to retrieve aerosol properties at night-time with the GRASP algorithm. Atmospheric Chemistry and Physics, 19(22), 14149-14171.

Herreras, M., Román, R., Cazorla, A., Toledano, C., Lyamani, H., Torres, B., ... & de Frutos, A. M. (2019). Evaluation of retrieved aerosol extinction profiles using as reference the aerosol optical depth differences between various heights. Atmospheric Research, 230, 104625.

Page 6, lines 183-185: The authors introduce here a past comparison between Pandora and Brewer without giving any result of this comparison. This sentence seems dispensable if it does not provide more information about the validity of Pandora NO_2 data.

Comparison results have been included in the text as follows:

"Total NO_2 data from the Pandora instrument #117 located at APL-SAP have been compared with NO_2 observations retrieved by the co-located MkIV Brewer spectrophotometer with serial number #067, revealing a correlation coefficient above 0.96 and a negligible absolute median bias of 0.002 DU (Diémoz et al., 2021)."

Page 6, last paragraph: This paragraph explains the NO_2 deviation Pandora versus OMI (AERONET) as is displayed in Fig. 4. It is written that, according to Fig. 4 (lower panel), biases (Pandora-OMI, I guess) of 89% and 87% are found. I'm not able to see these results in the lower panel. Later, the authors give another result: Pandora-OMI average differences of 61.5% at both stations. Could you please explain more in detail these different results and where do they come from?

The distribution of Pandora-OMI deviations are shown in the lower panel of Fig. 4. It should be noted that Fig. 4 have been changed, but the lower panel still shows the same distributions as before, without stacked bars. The peaks of the distribution are between 0-0.2 DU and the calculated mean biases (with the respective percentage values) are 0.15 ± 0.19 DU (61.5 ± 71.5%) and 0.16 ± 0.18 DU (61.5 ± 67.2%) for APL-SAP and CNR-ISAC, respectively. The 89% and 87% values are not the estimated biases. These numbers indicate the percentage of the Pandora-OMI data pairs with positive deviations, i.e. for how many cases the Pandora values are higher compared to the respective OMI values. Since the initial text might be misleading, it has been revised.

Page 9, line 281: As mentioned before, the authors acknowledge throughout the text that this database is too short to perform statistically meaningful trend analysis. The question is obvious: why then carry out this analysis?

The purpose of the analysis is not to derive any conclusions on the AOD or AE trends in the two Rome sites. The sole purpose of showing the trend analysis is to quantify how much the trend differs for the standard and the modified AODs. The importance of this difference can be assessed only quantitatively as aspects such as the trend analysis regression uncertainty and the trend “levels” compared with actual AOD uncertainties are also affecting the climatology related results. However, for a location with a significant NO_2 trend, the AOD trend will be also affected. The Rome AOD climatology and trends is a topic of a follow up, in progress study.

Page 11, lines 311-314: A reference to previous studies in Rome including some climatological data and aerosol types would be useful in this context.

As mentioned above, the climatology of Rome is not the main goal of the study. Some relevant references have been included in the text, which is modified as follows:

“Interestingly, based on Fig. 6, the highest Pandora NO_2 retrievals (reddish colors) are not associated with the highest AOD values, indicating that in Rome the high AOD loadings are not strictly associated with high NO_2 pollution events. In fact, high AODs are frequently related to long-range transport of elevated layers of desert dust, fires plumes or a combination of both (e.g., Barnaba et al., 2011; Gobbi et al., 2019; Campanelli et al., 2021; Andrés Hernandez et al., 2022). Hence, it might be worth to modify aerosol retrievals for high NO_2 in those pollution-related events with low to medium AOD levels. More about AOD and aerosol type climatology for the Rome area can be found in Di Ianni et al., (2018) and in Campanelli et al. (2022). “

Andrés Hernández, M. D. et al.: Overview: On the transport and transformation of pollutants in the outflow of major population centres – observational data from the EMeRGe European intensive operational period in summer 2017, Atmos. Chem. Phys., 22, 5877–5924, https://doi.org/10.5194/acp-22-5877-2022, 2022.

Barnaba, F., Angelini, F., Curci, G., and Gobbi, G. P.: An important fingerprint of wildfires on the European aerosol load, Atmos. Chem. Phys., 11, 10487–10501, 10.5194/acp-11-10487-2011, 2011.

Campanelli, M., Iannarelli, A.M., Mevi, G., Casadio, S., Diémoz, H., Finardi, S., Dinoi, A., Castelli, E., di Sarra, A., Di Bernardino, A., Casasanta, G., Bassani, C., Siani, A.M., Cacciani, M., Barnaba, F., Di Liberto, L., Argentini, S.: A wide-ranging investigation of the COVID-19 lockdown effects on the atmospheric composition in various Italian urban sites (AER – LOCUS), Urban Climate, Volume 39, 100954, ISSN 2212-0955, https://doi.org/10.1016/j.uclim.2021.100954, 2021.

Campanelli, M., Diémoz, H., Siani, A. M., di Sarra, A., Iannarelli, A. M., Kudo, R., Fasano, G., Casasanta, G., Tofful, L., Cacciani, M., Sanò, P., and Dietrich, S.: Aerosol optical characteristics in the urban area of Rome, Italy, and their impact on the UV index, Atmos. Meas. Tech., 15, 1171–1183, https://doi.org/10.5194/amt-15-1171-2022, 2022.

Di Ianni A, Costabile F, Barnaba F, Di Liberto L, Weinhold K, Wiedensohler A, Struckmeier C, Drewnick F, Gobbi GP.: Black Carbon Aerosol in Rome (Italy): Inference of a Long-Term (2001–2017) Record and Related Trends from AERONET Sun-Photometry Data. Atmosphere. 9(3), 81, https://doi.org/10.3390/atmos9030081, 2018.

Gobbi, G.P., Barnaba, F., Di Liberto, L., Bolignano, A., Lucarelli, F., Nava, S., Perrino, C., Pietrodangelo, A., Basart, S., Costabile, F., Dionisi, D., Rizza, U., Canepari, S., Sozzi, R., Morelli, M., Manigrasso, M., Drewnick, F., Struckmeier, C., Poenitz, K., Wille, H.: An inclusive view of Saharan dust advections to Italy and the Central Mediterranean, Atmospheric Environment, 201, 242-256, 10.1016/j.atmosenv.2019.01.002, 2019.

Page 11, line 332: The values of 1.1% and 1.9% included in this line (as well as in the following lines) don't correspond to the values in the table. Are the authors reducing the floating points in the text? The use of these values can cause confusion in the reader.

The authors thank the referee for noticing that. Indeed, we reduce the floating points in the text. Since a precision higher than one decimal point is not necessary for the percentage values, tables have been revised accordingly. The way the numbers are presented in the text has been changed in order to be consistent with the table.

Page 12, lines 350-352: The authors stated that, according to Table 2, SKYNET retrievals are quite similar irrespective of the TROPOMI data used. However, similar results (low difference with the PNG product) were retrieved also in Table 1 for Pandora. Furthermore, mean deviations of AERONET products also displayed very low values…

Indeed, the mean deviations of AERONET products are very low. However, these values decrease by about a half when the monthly TROPOMI NO_2 is used. That's the case for all AERONET products (AOD at 380/440nm and AE). One would expect similar behavior for SKYNET retrievals, which is not the case. This comment refers only to the behavior observed using TROPOMI NO_2. In the results presented in Table 1 there is not such a clear pattern of decrease when monthly values are used, even for AERONET products.

Page 12, Figs. 8 and 9: Why not merge these two figures into one?

The figures have been merged into one (Fig. 8).

Page 12, line 363: Please define what "modified AOD" is.

The word "modified" has been changed to "NO_2-modified" and a reference to Sections 2.4.1 and 2.4.2 has been added.

Page 12, lines 365-370: I find relevant the lack of information (numbers) to quantify these results.

Numbers have been added in the text as follows:

"The median AOD bias for AERONET is about 0.003 with a maximum of about 0.02 at the peak of the event. The median and maximum AE biases are 0.014 and 0.11, respectively. It can be also noted that in the case of SKYNET both AOD (median value of ~0.008 with a maximum of ~0.03) and AE deviations

(median and maximum values of ~0.03 and 0.10, respectively) are a bit higher compared to the respective AERONET deviations of synchronous data."

Page 14, section 3.6: The authors stated in section 2.2.1 that level 1.5 SSA AERONET data were used in this paper. However, in this section, it is not clear to me what SSA product was used. If I understand well, a mimic of the AERONET product retrieved by GRASP was used as a reference, instead of the AERONET SSA standard product. If so:

- Please correct the information provided in section 2.2.1 including a suitable explanation of GRASP algorithm and products used in this paper.
- Why not use the real product instead a "mimic" product? At least these two SSA should be compared…

Could you please clarify it?

The paragraph about SSA has been removed from section 2.2.1 and a new section (Sect. 2.6) has been included to clearly explain GRASP algorithm and the two approaches selected for this study.

Additionally, in Section 3.6 specific clarification has been added to explain why the GRASP algorithm has been used for the proposed comparisons instead of the AERONET product. The GRASP and AERONET inversion algorithms are fundamentally very similar. GRASP was borne from the heritage of AERONET. However, the different developments of both codes now imply some differences in the provided retrieval products. Thus, to avoid any source of discrepancy that is not introduced purely by the methodology to account NO_2, the authors consider that the most appropriate comparison should be done with two identical applications of GRASP, but with different NO_2 information.

Comprehensive and meaningful comparisons of the GRASP and AERONET retrievals is a very interesting topic. However, it would need specific investigation, which is out of the scope of this study.

Page 15, Fig. 12: There is no information about the lower panel plot. Is the SSA difference?

The lower panels represent the probability density functions of the differences in SSA at 440 nm between both GRASP approaches in each of the selected stations (Fig. 11 in the revised manuscript). The corresponding explanations and corrected labels have been added in the manuscript.

Page 15, Fig. 12: Y-axis of the upper plot should be SSA and not NO_2.

The labels in the axis of this figure represent the name of the applied methodology (Fig. 11 in the revised manuscript). All points shown there correspond to SSA at 440 nm. Corresponding description has been included in the manuscript.

Page 15, Fig. 12: Information about correlation is written in the text in terms of r-squared while in this figure is expressed as correlation coefficient "R" (in capital letters). Please homogenize.

The text has been revised to be consistent with the figure (Fig. 11 in the revised manuscript).

Page 15, line 454: Again, the numbers provided in the text do not correspond to the ones provided in the plot. It is a matter of rounding correctly to the appropriate significant digit. For example: with RMSE values of 0.035 and 0.031 I don't consider it appropriate to conclude that RMSE is < 0.035. The same for R squared.

A correct rounding has been applied to the corresponding numbers in the text.

Page 15, line 452: Why the threshold of 0.9 DU?

The comparisons and text have been updated with a new threshold of 0.7 DU, which corresponds to the average NO_2 concentration of the whole data set (0.4) plus 2 times the standard deviation. This change has been made to provide statistical significance to the selected value.

Page 15, line 453-454: The authors stated that a positive bias of 0.02 was found in conditions of high NO_2 concentrations. Are they talking about SSA or NO_2? From what figure (upper or lower panel) this result comes from? I see in the lower panel an average difference of 0.01 for $NO_2 > 0.9$ (high NO_2 conditions) but 0.02 for all conditions. From where did the authors find this result? I feel lost with this section.

The clarifications in the text and the correct labels of the lower panels in Fig. 11 that have been added to the manuscript now clearly show from what data these conclusions have been reached. All data in Fig. 11 correspond to SSA at 440 nm. The probability density functions representation show a consistent average of the difference between both methodologies of 0.03 for values of SSA lower than 0.9 and a mean difference of 0.01 for values of SSA higher than 0.9.

Page 15, lines 455-458: This sentence seems confusing to the reader. Please rephrase. It has also some typos, like the comma after the word "studies".

The corresponding sentence has been revised to improve clarity as follows:

"Previous studies have found SSA retrieval uncertainties in the range of 0.02-0.03 (Eck et al., 2003; Corr et al., 2009; Jethva et al., 2014; Kazadzis et al., 2016), whereas the correction, when high NO_2 is recorded, is usually higher."

Page 16, line 477: The general result stated here (AOD differences below 0.01 because of this NO_2 correction) seems really relevant. In fact, this is the main result a reader is expecting. However, is this general result written somewhere in the text?

Results have been added in the discussion of Section 3.1 as follows:

"The estimated AOD and AE deviations are below 0.01 and 0.1, respectively, for the majority of observations, i.e., about 96 - 98% of occurrences for both CNR-ISAC and APL-SAP (see also distributions in Fig. 6). The average AOD bias is between 0.002 ± 0.003 and 0.003 ± 0.003 (with the higher values observed at 380nm), while the average AE bias is ~0.02 ± 0.03. Overall, the mean AOD bias is low compared to the estimated uncertainties for the standard AERONET product, i.e., 0.01 - 0.02 (with the higher errors observed in the UV) (Sinyuk et al., 2020). However, the mean AOD bias for the cases of high NO_2 levels (> ~0.7 DU) is ~0.011 ± 0.003 at 440 nm and ~0.012 ± 0.003 at 380 nm for APL-SAP and ~0.009 ± 0.003 at 440 nm and ~0.010 ± 0.003 at 380 nm for CNR-ISAC, which is comparable to the AERONET reported uncertainties. The estimated mean bias of AE retrievals for the cases with high NO_2 (> ~0.7 DU) is ~0.08 ± 0.04 for both Rome sites. The threshold for NO_2 has been selected as the average Pandora NO_2 (~0.4) calculated from the whole data set plus two times the standard deviation...

… Similarly to AERONET, the derived AOD and AE biases for SKYNET are below 0.01 and 0.1, respectively, for the majority of observations, i.e., about 85% of occurrences for AOD and about 90% for AE (see also distributions in Fig. 7). The overall average AOD bias is ~0.007 ± 0.003, which can be assumed low considering that Nakajima et al. (2020) have estimated a root-mean-square difference (RMSD) of about 0.03 for wavelengths < 500 nm in city areas in AOD comparisons with other networks. However, the mean AOD bias for the cases with high NO_2 levels (> ~0.7 DU) is found to be about 0.018 ± 0.003, which is comparable to the RMSD value reported by Nakajima et al. (2020). The overall average AE bias calculated in this study is ~0.05 ± 0.04, whereas the AE bias averaged over the high NO_2 cases is about 0.10 ± 0.05."