Response to anonymous referee #2

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).

Reviewers' comments:

The authors use extensive measurements to investigate the impact of NO_2 concentrations on AOD and AE retrievals. This paper contributes to better understanding that considering NO_2, which is highly diurnal-variable, is important to improve aerosol properties in the spectral range where NO_2 absorption is strong. Since the manuscript is well-written, I think readers may understand your approach and result well. I believe the paper can be published for AMT after addressing the concerned expressed below.

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 Issues and specific comments:

P4 L104:

In AERONET site information (https://aeronet.gsfc.nasa.gov/new_web/photo_db_v3/Rome_Tor_Vergata.html),

Rome-Tor Vergata site is located at elevation=130 m but your description is shown as 117 m.

Which one is correct?

117 m is the value given in PGN data files and corresponds to the altitude at ground level. The value of 130 m in AERONET site info refers to the elevation on the roof where the instrument is installed.

P6 L166

Do you use NO_2 VCD (vertical column density) or SCD (slant column density) from Pandora product? For clarification, it might be better to mention you use NO_2 VCD in Section 2.3.1

We used the vertical column of NO_2. The text has been revised accordingly.

P6 L183:

Do Brewer NO_2 and Pandora NO_2 show good agreement? You need to mention how good quality in your Pandora NO_2 measurement since you use Pandora NO_2 to correct AERONET and SKYNET operational AOD, AE, and SSA product. More reliable NO_2 measurements make your study more meaningful. So, add one or two sentences to show how Pandora NO_2 agrees well with NO_2 from other instruments.

Results from the comparisons with Brewer and MFDOAS NO_2 as well as estimations of the Pandora total NO_2 accuracy have been included in the text:

"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). According to Herman et al. (2009), the Pandora direct-sun total NO_2 has a clear-sky precision of 0.01 DU in slant column and a nominal estimated accuracy of 0.1 DU in the vertical column. In the same study, a systematic difference of less than 1% was found between the relative slant columns of Pandora and a MultiFunction Differential Optical Absorption Spectroscopy (MFDOAS) instrument."

P6 L191: The Pandora data -> The Pandora NO_2 data

The text has been revised accordingly.

P7 L198: You did not show the absolute NO_2 difference. However, I think Pandora NO_2 is one of the most essential parts in your method. So, it had better to create this plot in the main or the supplement to show how much absolute difference between Pandora NO_2 and climatology OMI. If so, readers will understand your approach better.

The difference between Pandora NO_2 and OMI climatology is discussed in the text by presenting both absolute and percentage mean values with standard deviation:

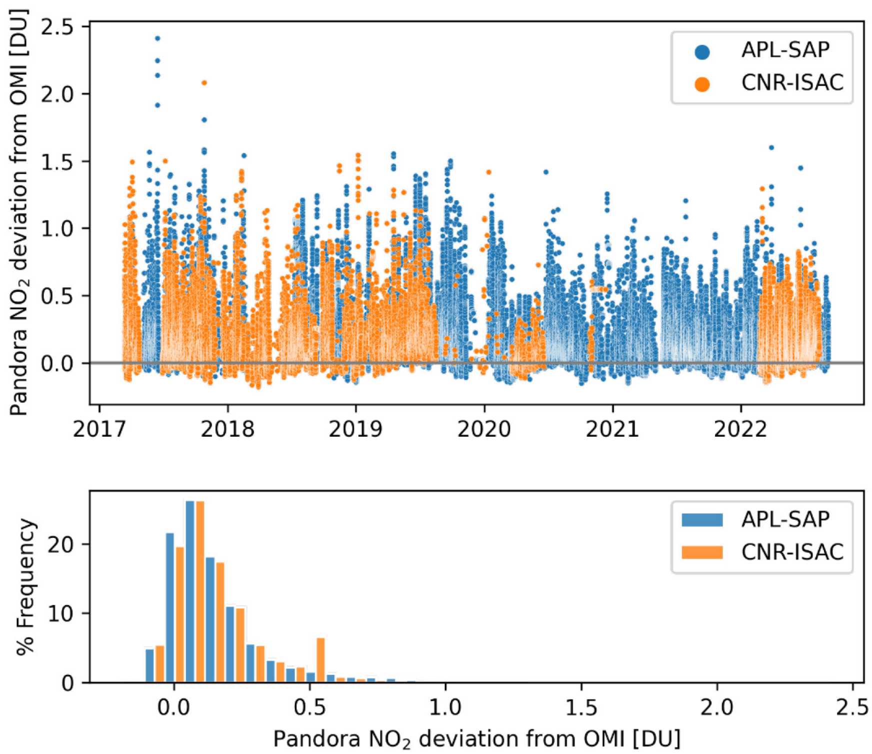
"AERONET aerosol retrievals seem to significantly underestimate the NO_2 abundance over urban and suburban locations with an average absolute difference between the actual Pandora measurements and the estimations from satellite climatology of about 0.15 ± 0.19 DU (61.5 ± 71.5%) and 0.16 ± 0.18 DU (61.5 ± 67.2%) for APL-SAP and CNR-ISAC, respectively."

We also discuss the range of the derived biases (both absolute and percentage) within which the majority of cases is observed:

"The majority of PGN-OMI biases lie within 0-0.5 DU corresponding to Pandora values lower than 1 DU. More specifically, 90% of the PGN NO_2 data over APL-SAP differ within -0.14 DU (-50%) and 0.44 DU (150%) from OMI climatology, while the respective deviation range between -0.14 and 0.51 DU (-50% – 170%) for CNR-ISAC. However, there are quite a few cases (~9.5% and ~8.8% for APL-SAP and CNR-ISAC, respectively) of higher PGN values (< 2 DU) leading to larger deviations (up to ~1.6 DU for APL-SAP and ~1.5 DU for CNR-ISAC)."

Thus, we think that the levels of the absolute difference between Pandora NO_2 and OMI climatology, as well as their distribution, are clearly presented by the numbers included in the text.

However, based on this comment and a comment from referee #3, the upper panels of Fig. 4 have been replaced with the time series of Pandora – OMI deviations (see figure below).



P7 P225: Are there any specific reasons to exclude the COVID-19 lockdown period? If so, please mention briefly.

Since the TROPOMI data cover a relatively short period (2018-2021) and Fig. 5 is for visualization purposes only, we excluded the lockdown period in order to prevent the low values observed during that period from affecting the average NO_2 values. A brief explanation has been included in the revised manuscript.

P9 L262: In AERONET (Eck et al., 1999), AE is -> The AERONET AE product (Eck et al., 1999) is

The text has been revised accordingly.

P9 L279: the impact of AOD and AE modified retrievals -> the impact of modified AOD and AE retrievals

The text has been revised accordingly.

P9 L280: to investigate the possible effect on the AOD and AE trends -> to investigate the possible effect of NO_2 absorption on the AOD and AE trends

The text has been revised accordingly.

P10 L299: to investigate the impact of AOD and AE modified calculations on the derived temporal trends -> to investigate the impact of modified AOD and AE calculations on the derived temporal trends

The text has been revised accordingly.

P11 L311: Any references? Or is this your finding in this research? Then, plot it to explain or direct the figure you show this. You can show the correlation between NO_2 and AOD.

This is a finding from Fig. 6. Reddish colors (indicating high NO_2 values) do not correspond to high AOD loadings. A reference to the figure has been added in the text.

P11 L336: Do you have any reason to use SKYNET AE for 400-1020 nm?

You use AERONET AE for 440-870 nm. Then, is it more consistent to use similar wavelength pair like SKYNET AE for 400-870 nm?

The aim of this study was to investigate the impact of NO_2 absorption on the standard network products, i.e., the products officially available online. The only AE product available from SKYNET is at wavelengths 400-1020nm. In addition, there is not any AE product from AERONET at a spectral range closer to 400-1020nm than AE at 440-870nm.

P11 L338: You show how modified AOD and AE by considering Pandora NO_2 and then show modified AOD and AE by implementing TROPOMI NO_2. Reader can ask how Pandora NO_2 and TROPOMI NO_2 are consistent. It had better to add one or two sentences to show how both NO_2 measurements are in good agreement. You can refer previous studies about this.

References of TROPOMI and Pandora total NO_2 comparison studies have been included in the manuscript:

"Despite the improved spatial resolution of TROPOMI, the NO_2 corrections using TROPOMI data are expected to be less accurate than those performed with the Pandora product. For example, Lambert et al. (2021) showed a bias between TROPOMI and Pandora total NO_2 column ranging from -23% over polluted stations to +4.1% over clean areas with a median bias of -7.1%, in the frame of the standard validation process of TROPOMI Level 2 NO_2 products. Other studies have concluded similar results. For example, Zhao et al. (2020) showed a negative bias for the standard TROPOMI total NO_2 product in the range 23 - 28% over urban and suburban environments and a positive bias of 8 - 11% at a rural site,

while Park et al. (2022) showed 26 - 29% negative bias and R^2 within 0.73-0.76 over the Seoul Metropolitan Area in Korea."

P13 L381: The results -> The results in Table 3

The text has been revised accordingly.

P13 L381-388: The description is the analysis in Table 3. Readers may also be curious about the trend itself. AE trends in CNR-ISAC and APL-SAP shows positive and negative, respectively. Do you have any interpretation for this? Is it because inhomogeneous local emission patterns and photochemical destruction you mentioned in P15 L465? Or during your trend analysis period, were there more frequent transports of dust from Africa and caused it negative AE trend in APL-SAP?

We think that it is not possible to answer to this question without speculating based on the absolute changes of the AE, the limited (for such analysis and interpretation) period, the various sources of AE trend uncertainty and the fact that the two datasets are not directly comparable since they are not synchronous. To elaborate a bit more, AE trends for both stations end up in the same range with the AE retrieval uncertainty based on the AOD uncertainty and also comparable with the standard error of the regression slope.

Based on the above reasons, we tried to avoid to present that analysis as a climatology of the area and just used it as an assessment of the error propagation of NO_2 correction to AOD and AE trends.

The positive AE trend in the limited time period addressed in this work is just a short-term effect, not a long-term one. A long-term analysis with the CNR-ISAC data (> 20 years) is in progress and will be presented in a follow-up investigation. We can anticipate that, at that site, there is a clear negative trend of fine-fraction AOD, while coarse-AOD keeps almost constant, and this translates into a decreasing AE.

P13 L402: Font type looks different.

The font type has been corrected.

P14 L432: You used not standard AERONET aerosol retrieval but GRASP algorithm.

If both are the same condition, retrieved SSA from GRASP algorithm is the same with that from standard AERONET retrieval? If not, how much difference of SSA is apparent?

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 and necessary study. However, the level of required detail and deepness is totally out of the scope of this study.

Also, when you use SSA from AERONET, there are quality assurance criteria (Mok et al., 2018). In Figure 12, do you plot SSA when AOD > 0.4? SSA when AOD is small shows large error.

The authors agree with the referee in the conditions established by Mok et al. (2018) as well as by Sinyuk et al. (2020). 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 loose of accuracy in the retrieved values results beneficially for the final quality of the comparisons. 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) and Herreras et al. (2019).

However, despite these uncertainties, in the methodology proposed here the comparison is made with identical retrieval schemes but with different NO_2 representation. Thus, even if random error is present in the retrieved values of SSA at 440 nm, the error observed here is a systematic bias. This is why the conclusions about the need of a correct representation of this gaseous absorption in AERONET-like retrievals are not affected by possible inconsistencies in the amount of information available under AOD or Solar Zenith Angle conditions.

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.

In addition, for SSA calculation, I am wondering you use the consistent surface albedo for SSA retrievals. Incorrect surface albedo makes a systematic bias in SSA retrievals (Mok et al., 2018).

Mok, J., Krotkov, N. A., Torres, O., Jethva, H., Li, Z., Kim, J., Koo, J.-H., Go, S., Irie, H., Labow, G., Eck, T. F., Holben, B. N., Herman, J., Loughman, R. P., Spinei, E., Lee, S. S., Khatri, P., and Campanelli, M.: Comparisons of spectral aerosol single scattering albedo in Seoul, South Korea, Atmos. Meas. Tech., 11, 2295–2311, https://doi.org/10.5194/amt-11-2295-2018, 2018.

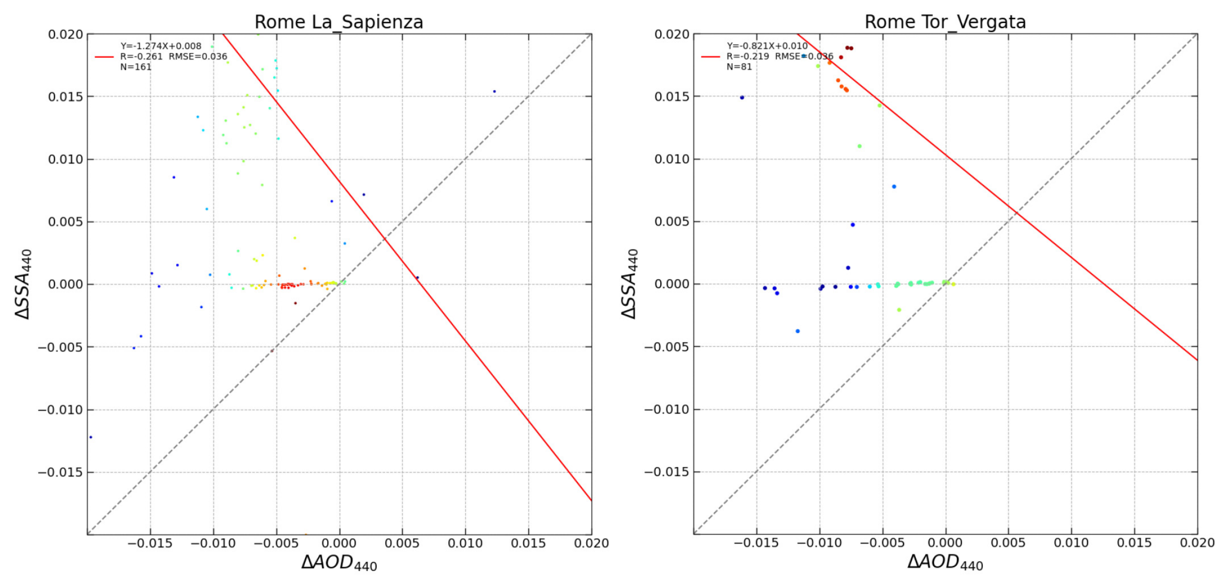
Both GRASP retrieval schemes are applied using the BRDF as described in Román et al. (2018), a bi-weekly climatology of MODIS BRDF product over the corresponding AERONET sites.

Lastly, overestimation in AOD lead to the underestimation in SSA. When you compare SSA from GRASP/Standard AERONET with that from GRASP/Pandora NO_2, do you use the same AOD?

For this, in figure 12, you should add a plot of difference of SSA as a function of difference of AOD.

The authors agree that AOD and SSA tends to be inversely correlated, and actually for the retrievals included in the figure (Fig. 11 in the revised manuscript) there is an underestimation of AOD represented by a MBE of -0.0068 (-4.5%). However, the retrieval is complex and a lot of parameters are part of the fitting. Thus, in order to fit TOD and almucantar the size distribution or particle sphericity can be affected by these different NO_2 conditions, which makes the direct connection between biases in AOD and SSA more complex.

The correlation between the differences in both magnitudes can be found below for both stations:



As it can be seen for AOD differences of less than 0.005, the differences in SSA remains very close to zero. However, establishing a direct relationship between both magnitudes require a deeper look to all parameters used to model aerosol particles.

P15 L458 or in conclusion:

You may add one or two sentences about the importance of your research to estimate the effect of NO_2 on the spectral dependence of SSA (i.e., absorption Ångström exponent (AAE)) as a future study.

The following text has been included in the conclusions:

"In future studies, the effect of NO_2 correction on absorption Ångström exponent (AAE) could be explored. AAE is an aerosol optical property that describes the absorption variation with respect to wavelength and is significantly influenced by particle size, shape, and chemical composition used for aerosol characterization and apportionment studies (e.g., Schuster et al., 2006). Since AAE is a function

of spectral AOD and spectral SSA, the NO_2 correction for certain AOD wavelengths and SSAs shown in this study is expected to impact the AAE calculations."

Schuster, G. L., Dubovik, O., and Holben, B. N.: Angstrom exponent and bimodal aerosol size distributions, J. Geophys. Res., 111, D07207, https://doi.org/10.1029/2005JD006328, 2006.

P27 L800 (Table 1)

Why NO_2 values in Table 1 is different in different wavelengths? Is this because the number of data you used for 380 and 440 is different? Why don't you use the same number of data at all wavelengths? Since we look at AE which is the relationship of AOD between wavelengths, I think you should match the data for all wavelengths. In case one event has some information at one wavelength is missing, it is caused by some issues like small fraction of cloud is passing etc.

The NO_2 quantity available from AERONET is the NO_2 optical depth. The NO_2 column in DU has been derived from the AERONET optical depth values using appropriate cross sections for each wavelength. Slight differences are observed between different wavelengths in the calculated columns (negligible differences in the third decimal point in DU) due to possible minor biases in the cross section used.

Another reason for the different deviations for different wavelengths presented in Tables 1 and 2 is, indeed, the different number of data. However, these differences in the number of data are not due to sparse events, but they refer to whole periods during which the Cimel instrument would not perform measurements in one wavelength or another for any reason.

Cases with missing measurements for one or another wavelength were excluded from AE calculations.

P41 L890 (Figure 12)

In upper left figure, the number of data shown in the figure is not the same with the legend (N=32). Also, there is no explanations for different color (e.g., green and red dots). It is hard to recognize the dots in the plot. Please modify them with increasing size.

The number of points in the plot is correct. The color is an indicator of the density of points, i.e., colors closer to red indicate higher amount of points close together. A very high density of cases with SSA values > 0.95 is observed and this is why it is difficult to distinguish by eye the total number of points stated in the legend. The size of the points has been increased and explanation for the different colors has been added (Fig. 11 in the revised manuscript).