Thank you for your careful review and constructive suggestions. These suggestions are quite valuable to us, and help improve our manuscript a lot.

Point-to-point responses

We appreciate the reviewers for their valuable and constructive comments, which are very helpful for the improvement of the manuscript. We have revised the manuscript carefully according to the reviewers' comments. We have addressed the reviewers' comments on a point-to-point basis as below for consideration, where the reviewers' comments are cited in ***black****, and the responses are in* ***blue****.*

I'm pleased to accept your revised manuscript "Evaluation of Transport Processes over North China Plain and Yangtze River Delta using MAX-DOAS Observations" for publication in ACP subject to minor revisions as listed below.

As the reviewer pointed out, your explanation of the Ring effect is not correct.
Re: Thank you for this comment.
"The Ring spectrum was added to the fitting settings to remove the influence of the stratosphere on the DSCDs." -> "The Ring spectrum was added to the fitting settings to remove the influence of inelastic rotational Raman scattering on solar Fraunhofer lines (Chance and Spurr, 1997; Grainger and Ring, 1962)."

Please check again your error discussion - I agree with the reviewer, that it could be improved
Re: Thank you for this comment. We have changed our demonstrations in Section 2.4 and Supplementary Sect. S3 as follows.
Section 2.4: "Algorithm error (i.e., the difference between the measured and modeled DSCDs) mainly arises from an imperfect representation of the real radiation field in the RTM - spatial inhomogeneities of absorbers and aerosols, clouds, real aerosol phase functions etc."
Supplementary Sect. S3: "Algorithm error is the discrepancy between the measured ($\boldsymbol{y}$) and modelled DSCDs ($F(\boldsymbol{x}, \boldsymbol{b})$). As displayed in Eq. s3, the error sources that result in this discrepancy include forward model error from an imperfect approximation of forward function F, forward model parameter error from selection of parameters $\boldsymbol{b}$, and errors not related to the forward function parameters, like detector noise (Rodgers, 2004). Algorithm error is a function of the viewing angle. Due to the difficulty of assigning this error to each altitude of profile, the algorithm errors on the near-surface values and column densities are usually estimated by calculating the average relative differences between the measured and modeled DSCDs at the minimum and maximum elevation angle (except 90°), respectively (Wagner et al., 2004).

$$\sigma_{algorithm} = \mathbf{y} - F(\mathbf{x}, \mathbf{b}) \quad \text{(s3)}$$

where $F(\boldsymbol{x}, \boldsymbol{b})$ is the forward model; $\boldsymbol{b}$ represents the meteorological parameters; $\boldsymbol{y}$ is the measured DSCDs; $\boldsymbol{x}$ is the state vector."

In the abstract (line 25), please add the region (China or more specific) your measurements took place.
Re: Thank you for this comment.
"… and analyzed three typical transport phenomena." -> "… and analyzed three typical transport phenomena over the North China Plain (NCP) and Yangtze River Delta (YRD)."

Page 4, line 3 please remove "technological"
Re: "…, providing technological support for horizontal pollutant transport analysis." -> ", providing support for horizontal pollutant transport analysis."

Page 4, line 95: While it is correct that MAX-DOAS does not require absolute radiometric calibration, to my knowledge, the other techniques mentioned also don't require it. This is not the main advantage of MAX-DOAS.
Re: "Compared with the above techniques, MAX-DOAS does not require radiometric calibration and has many other advantages such as simple design, low power demand, possible automation, low cost, and minimal maintenance." -> "Compared with the above techniques, MAX-DOAS has many advantages such as simple design, low power demand, possible automation, low cost, and minimal maintenance."

Table 2: Please correct the units for O4

Re: "293 K, I_0 correction (SCD of 3×10^{43} molecules cm^{-2}); (Thalman and Volkamer, 2013)" -> "293 K, I_0 correction (SCD of 3×10^{43} $molecules^2$ cm^{-5}); (Thalman and Volkamer, 2013)"

Page 8, line 144: I do not think that this is what happens in your analysis. You cannot derive the SCDs and then take the difference - I assume that you derive the DSCD directly by using the zenith observation as the background I_0 in the DOAS equation.
Re: Thank you for this comment. We added this sentence to explain what SCDs are, but our demonstration here was inaccurate. We have corrected this mistake.
"Spectral analysis derives the slant column densities (SCDs), i.e., the integrated concentration along the light path. Subsequently, we calculated the differential slant column densities (DSCDs), which are defined as the difference between the off-zenith and zenith SCDs." -> "Slant column density (SCD) is defined as the integrated concentration along the light path. Firstly, we calculated the differential slant column densities (DSCDs), which are defined as the difference between the off-zenith and zenith SCDs. Subsequently, we analyzed the…"

Page 8, line 154: The explanation of the detection limit seems not correct

Re: Thank you for your comment. We have corrected this mistake and recalculated the detection limit.
"We assumed two times the fitting error RMS as the DSCD detection limits (Wang et al., 2017; Lampel et al., 2015), which were 7×10^{41} ($molec^2 \cdot cm^{-5}$), 1.6×10^{15}, 3.6×10^{15}, and 5.8×10^{14} $molec \cdot cm^{-2}$ for O_4, NO_2, HCHO, and HONO, respectively."

-> "The DSCD detection limits were roughly estimated using two times of the mean RMS divided by the absorption cross-section (Nasse et al., 2019; Wang et al., 2017; Lampel et al., 2015), which were 2.4×10^{42} ($molec^2 \cdot cm^{-5}$), 1.7×10^{15}, 8.9×10^{15}, and 2.5×10^{15} $molec \cdot cm^{-2}$ for O_4, NO_2, HCHO, and HONO, respectively."

Page 9, line 168: I assume that the RTM model covers the full atmosphere and only your inversion assumes that there is no aerosol / trace gas above 3 km? Is that a good assumption in case of transport events?
Re: Thank you for this comment. This suggestion reminds us that we have missed some retrieval information in our manuscript. We have added it in our manuscript as follows.
"The vertical distribution of trace gas above the retrieval height (3 km) was fixed to follow the U.S. Standard Atmosphere (Anderson et al., 1986)."
For some high concentration trace gases at high altitudes (e.g., O_3), their stratospheric profiles have been contained in RTM (Mayer and Kylling, 2005; Emde et al., 2016).

Page 9, line 170: In your description of the a priori, the surface concentration and the column are given, but not the assumed shape of the profile. Please add.
Re: Thank you for this comment. We have added the shape information of a priori profile in front of these sentences as follows.
"Exponentially decreasing functions with a scale height of 0.5, 0.5, 1.0, and 0.2 km were utilized as a priori profiles for aerosols, NO_2, HCHO, and HONO, respectively."

Page 9, line 190: I assume this difference is not from the "imperfect minimum of the cost function" but rather from the imperfect representation of the real radiation field in the RTM - spatial inhomogeneities of absorbers and aerosols, clouds, real aerosol phase functions etc.
Re: Thank you for this comment. We have changed our demonstration in Section 2.4 as follows.
"Algorithm error (i.e., the difference between the measured and modeled DSCDs) mainly arises from an imperfect representation of the real radiation field in the RTM - spatial inhomogeneities of absorbers and aerosols, clouds, real aerosol phase functions etc."

Overall flux discussion: In my opinion, the units you use are not a good choice. If you are interested in transport, then the number of molecules transported is relevant, not the

mixing ratio. The same mixing ratio at 3 km and at 0 km altitude lead to different numbers of transported molecules. Please convert to concentrations (molec / m3). This will also impact on your discussion of the most important transport levels.

Re: Thank you for this comment. We have supplemented the unit conversion process in Supplementary Sect. 4 as follows.

"To better demonstrate transport flux, we needed to convert trace gas mixing ratio (ppb) into molecular density (molec $\cdot$ m^{-3}) at first. The conversion formula involves temperature and pressure at different altitudes as follows.

$$C = \frac{X \bullet N_A}{V_m} \times 10^{-9} = \frac{X \bullet N_A \bullet P}{R \bullet T} \times 10^{-9} \quad \text{(s12)}$$

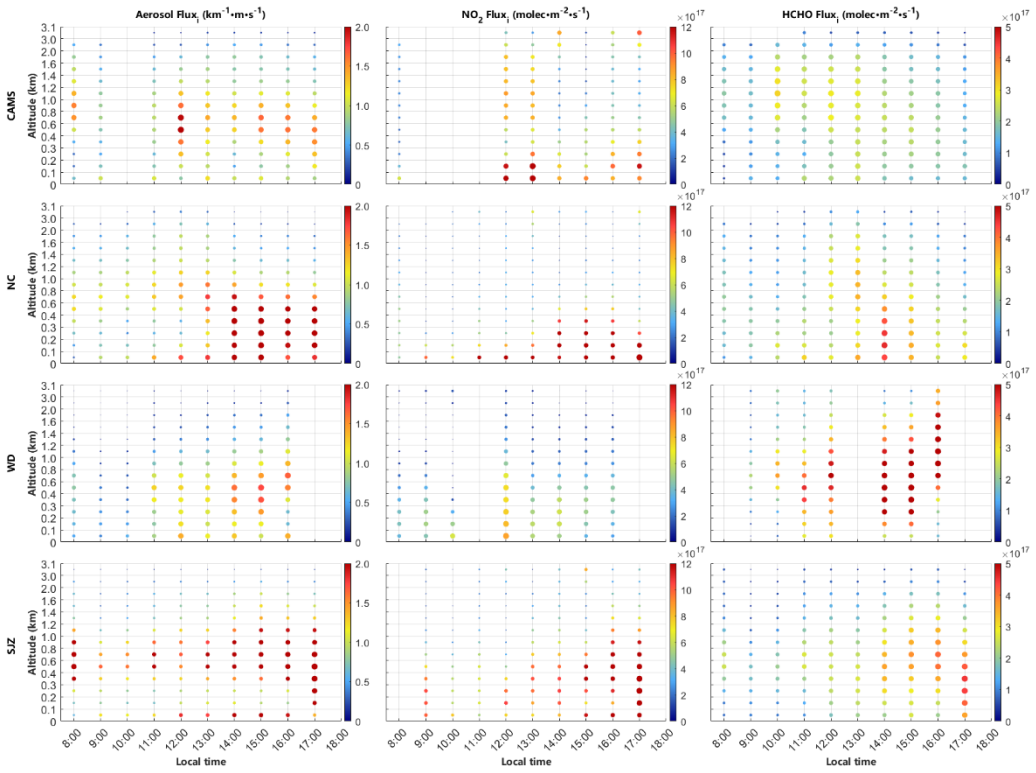
where C denotes the trace gas molecular density (molec·m^{-3}), and X is trace gas mixing ratio (ppb); N_A is Avogadro constant (6.02 × 10^{23} mol^{-1}); R is molar gas constant, with a value of 8.314 J·mol^{-1}·K^{-1}; P and T represent the atmospheric pressure and temperature at different altitudes, respectively. Berberan-Santos et al. (1997) described a relationship model which represents well the dependence of pressure and temperature on altitude for the whole troposphere (below 11 km) as follows.

$$T(z) = T_0 - \beta z \quad \text{(s13)}$$

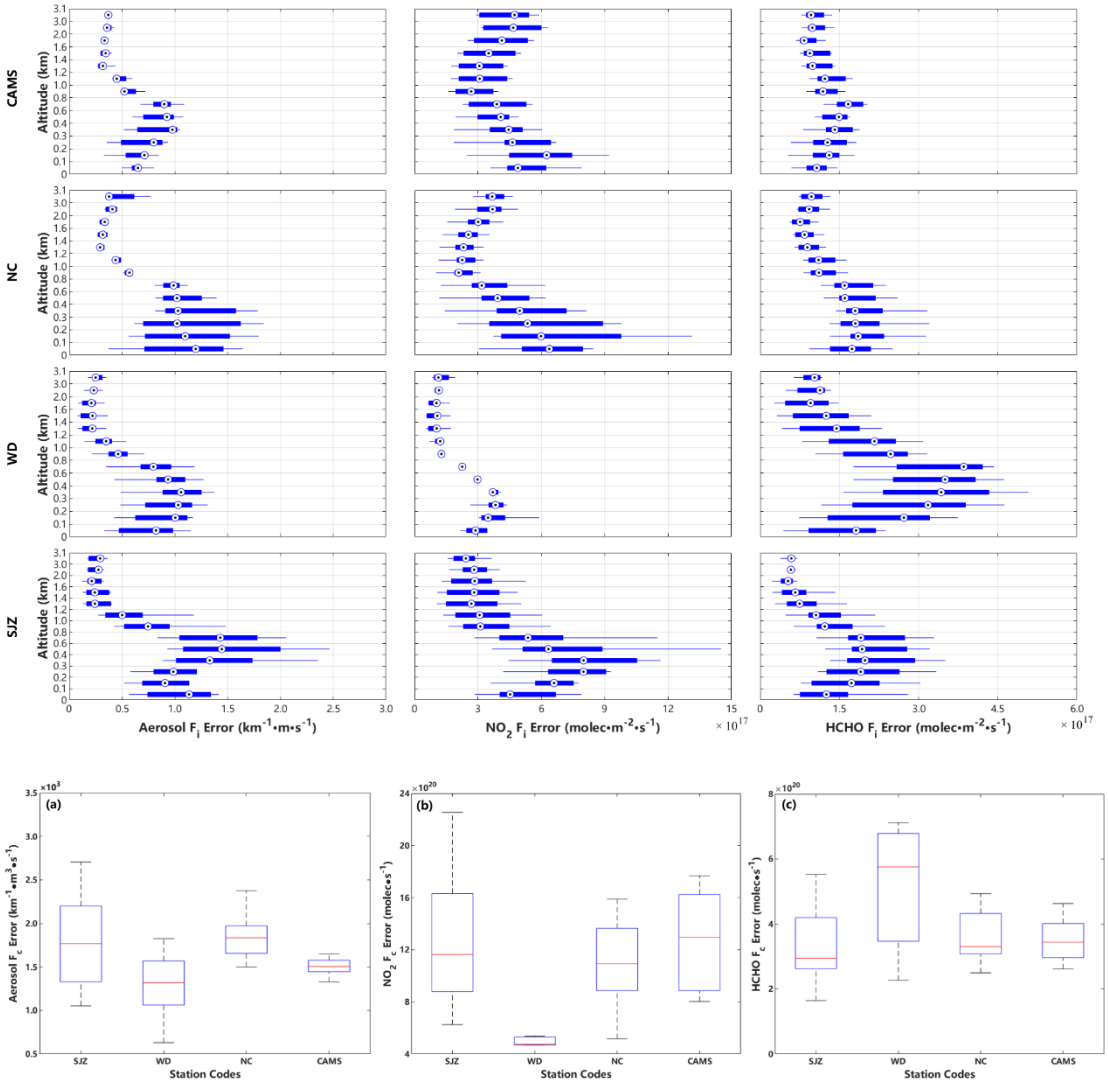
$$P(z) = P(0)\left(1 - \frac{\beta z}{T_0}\right)^{\frac{mg}{k\beta}} \quad \text{(s14)}$$

Here, T(z) and P(z) denote the temperature and atmospheric pressure at height z (km), respectively; T_0 and P(0) are the surface values; k is Boltzmann constant (1.38 × 10^{-23} J·K^{-1}); m is air molecular mass (29 × 10^{-3} kg·mol^{-1}); g represents acceleration of gravity (9.8 m·s^{-2}); β equals 6.5 K·km^{-1}."

Besides, we have re-depicted the transport flux variation figures (Fig. 4 and Fig. 5), and changed the transport flux descriptions in Section 3.1.

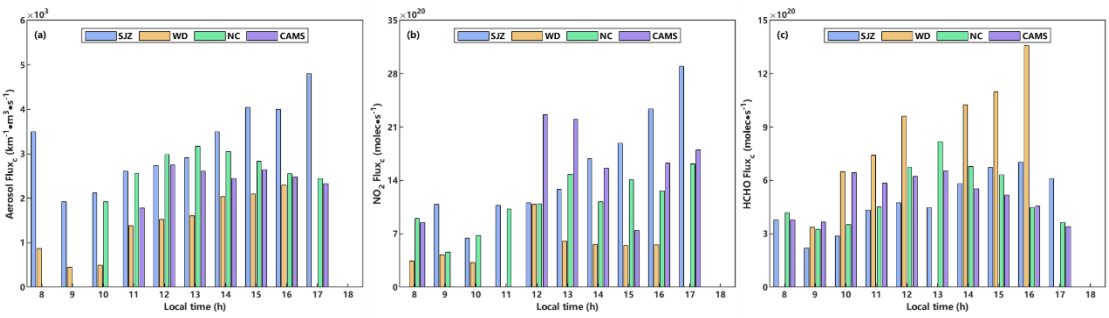


"Except that F_i reached the highest level of ~ 1.8 × 10^{18} molec·m^{-2}·s^{-1} in the 400–600 m layer at 16:00 at the SJZ station, the other highest F_i all occurred below 400 m at any station and at any time. This indicated that the MTL of NO_2 was 0–400 m. Near-surface NO_2 emission sources (e.g., vehicle and factory emissions) might be the main reason for this phenomenon. Compared with aerosols and NO_2, we found that high-value HCHO F_i extended to higher altitudes. Taking CAMS as an example, we found the strongest HCHO F_i constantly emerging at 1000–1200 m from 8:00 to 13:00, and averaging 2.51 × 10^{17} molec·m^{-2}·s^{-1}. During the same period, surface HCHO F_i only averaged 1.72 × 10^{17} molec·m^{-2}·s^{-1}. However, at the CAMS station, the surface HCHO concentration was much higher than that of the 1000–1200 m layer between 8:00 and 13:00 (Fig. 3), proving that high-altitude transport contributed more to overall HCHO transport. After 10:00, we found that the highest HCHO F_i gradually increased from ~ 3.5 × 10^{17} to ~ 4.5 × 10^{17} molec·m^{-2}·s^{-1} at WD, with the MTL of HCHO ranging from 400 to 1000 m. At station SJZ, the strongest HCHO F_i increased from ~ 2.6 × 10^{17} to ~ 4.5 × 10^{17} molec·m^{-2}·s^{-1} during 11:00–17:00, with the highest transport fluxes occurring mostly at 400–800 m. These findings indicated that the MTL of HCHO was mainly 400–1200 m."



Flux discussion: In addition to using concentrations, you should convert your numbers to real fluxes by integrating over a unit area of 1 x 1 m2. This will then result in a flux in units of molec / s through a unit area. In the caption of Figure 4, you claim that this is what you show, but the units you give to not match that claim.

Re: Thank you for this comment. We didn't use the $flux_i$ in units of molec / s, because the $flux_c$ was calculated by multiplying $flux_i$ and H_i, which would result in the $flux_c$ in units of $molec \cdot m \cdot s^{-1}$. This would result in a confusing physical meaning. After comprehensively considering, we planned to use the $flux_i$ in units of $molec \cdot m^{-2} \cdot s^{-1}$ and integrated the $flux_c$ over a unit width(1 m), making the $flux_c$ unit $molec \cdot s^{-1}$.



"For NO_2 transport, the average F_c values at SJZ (1.56×10^{21} molec $\cdot$ s^{-1}), NC (1.10×10^{21} molec·s^{-1}), and CAMS (1.58×10^{21} molec·s^{-1}) were substantially higher than those at WD (5.57×10^{20} molec·s^{-1}). Conversely, the average F_c of HCHO was the highest in WD (8.82×10^{20} molec·s^{-1}), whereas the F_c values in SJZ, NC, and CAMS were 4.81×10^{20}, 5.16×10^{20}, and 5.12×10^{20} molec·s^{-1}, respectively."

"Figure 4. Transport flux per unit cross-sectional area at different altitudes ($Flux_i$) at CAMS, NC, WD, and SJZ stations on February 5, 2021." -> "Figure 4. Transport flux at different altitudes ($Flux_i$) at CAMS, NC, WD, and SJZ stations on February 5, 2021."

Section 3.2: In the presence of a dust storm, the vertical sensitivity of the MAX-DOAS measurements will change. Have you considered that? It may explain some of your observations (disappearance of elevated layers).

Re: Thank you for this comment. It is difficult to exclude this possibility, because it is not easy to design the experiment to measure the effects of dust on vertical sensitivity. If placing the two identical instruments in two nearby places, the influence range of sand storm is so large that can cover the whole places. However, if the distances between two instruments are too far, the different environment would be introduced as another uncertain factor. In a word, it is hard to control variables and measure the effects of dust on the vertical sensitivity by doing experiments.

A previous study (using LiDAR) on dust storm also indicated that dust layers would inhibit the dissipation of pollutants and enhance surface air pollution, by depressing the PBL and weaken the turbulent exchange (Wang et al., 2020). This conclusion is in agreement with our findings. Thus, we thought the profile shape changes were more attributed to the accumulation of surface pollutants, and the effects of vertical sensitivity variation was relatively little compared to the actual pollutant increases.

However, we cannot completely exclude the possibility of the effects of dust on vertical sensitivity. Therefore, we have added it as one of uncertainties and demonstrated it at the end of Section 3.2.

"The comparison result between the dusty day and two clean days makes it possible to better understand the impacts of dust storm on local environment. However, there remain some uncertainties in this discussion. Although we selected the closest clean

days to lessen the effects of some factors (e.g., climate and temperature) on comparison, the uncertainties caused by other meteorological parameters (e.g., wind speed and directions) were unknown to us, since we did not make sure these parameters were nearly the same on these three days. Therefore, this comparison analysis is based on the assumption that there is little difference between meteorological parameters on various days or the effects caused by different meteorological parameters are negligible. Besides, a dust storm would trigger changes at the vertical sensitivity of MAX-DOAS measurements, which might influence profile shape. These impact factors are difficult to control in observations, and modelling correction may be a good solution."

Section 3.2: The discussion is based on three days - two "clean" days and one day during the dust storm. While your discussion is plausible, you should acknowledge that it is based on the assumption, that everything else is the same on the three days. This assumption is probably not correct - these are different days of the week (one is a Saturday), wind speed and directions are different and also the accumulation history of pollution in the BL is different. Very large day to day variations are observed at many stations without dust storms. Please at least mention this possibility.

Re: Thank you for this comment. This is one of uncertainties in our discussion, and we have added this uncertainty at the end of Section 3.2 as follows.

"The comparison result between the dusty day and two clean days makes it possible to better understand the impacts of dust storm on local environment. However, there remain some uncertainties in this discussion. Although we selected the closest clean days to lessen the effects of some factors (e.g., climate and temperature) on comparison, the uncertainties caused by other meteorological parameters (e.g., wind speed and directions) were unknown to us, since we did not make sure these parameters were nearly the same on these three days. Therefore, this comparison analysis is based on the assumption that there is little difference between meteorological parameters on various days or the effects caused by different meteorological parameters are negligible. Besides, a dust storm would trigger changes at the vertical sensitivity of MAX-DOAS measurements, which might influence profile shape. These impact factors are difficult to control in observations, and modelling correction may be a good solution."

References

Anderson, G. P., Clough, S. A., Kneizys, F., Chetwynd, J. H., and Shettle, E. P.: AFGL atmospheric constituent profiles (0.120 km), Tech. rep., AIR FORCE GEOPHYSICS LAB HANSCOM AFB MA, 1986.

Chance, K. and Spurr, R. J. D.: Ring effect studies: Rayleigh scattering, including molecular parameters for rotational Raman scattering and the Fraunhofer spectrum, Applied Optics, 36, 5224–5230, 10.1364/AO.36.005224, 1997.

Emde, C., Buras-Schnell, R., Kylling, A., Mayer, B., Gasteiger, J., Hamann, U., Kylling, J., Richter, B., Pause, C., Dowling, T., and Bugliaro, L.: The libRadtran software package for radiative transfer calculations (version 2.0.1), Geoscientific Model Development, 9, 1647-1672, 10.5194/gmd-9-1647-2016, 2016.

Grainger, J. F. and Ring, J.: Anomalous Fraunhofer line profiles, Nature, 193, 762,

10.1038/193762a0, 1962.
Mayer, B. and Kylling, A.: Technical note: The libRadtran software package for radiative transfer calculations – description and examples of use, Atmospheric Chemistry and Physics, 5, 1855-1877, 10.5194/acp-5-1855-2005, 2005.
Rodgers, C. D., Taylor, F. W. (Ed.): Inverse methods for atmospheric sounding, theory and practice, World Scientific, 255 pp.2004.
Wang, Z., Liu, C., Xie, Z., Hu, Q., Andreae, M. O., Dong, Y., Zhao, C., Liu, T., Zhu, Y., Liu, H., Xing, C., Tan, W., Ji, X., Lin, J., and Liu, J.: Elevated dust layers inhibit dissipation of heavy anthropogenic surface air pollution, Atmospheric Chemistry and Physics, 20, 14917-14932, 10.5194/acp-20-14917-2020, 2020.

Thank you for your careful review and constructive suggestions. These suggestions are quite valuable to us, and help improve our manuscript a lot.

Point-to-point responses

We appreciate the reviewers for their valuable and constructive comments, which are very helpful for the improvement of the manuscript. We have revised the manuscript carefully according to the reviewers' comments. We have addressed the reviewers' comments on a point-to-point basis as below for consideration, where the reviewers' comments are cited in ***black****, and the responses are in* ***blue****.*

The authors refer to the Ring effect as a kind of stratospheric correction. This is not correct. Instead, the Ring effect accounts for the filling-in of Fraunhofer lines caused by inelastic rotational Raman scattering [Grainger and Ring, 1962; Chance and Spurr, 1997]. This needs to be corrected.

Re: Thank you for this comment.

"The Ring spectrum was added to the fitting settings to remove the influence of the stratosphere on the DSCDs." -> "The Ring spectrum was added to the fitting settings to remove the influence of inelastic rotational Raman scattering on solar Fraunhofer lines (Chance and Spurr, 1997; Grainger and Ring, 1962)."

I feel that the definition of an "algorithmic error" is still not appropriate and remains vague. It is now mentioned that the "Algorithmic error (i.e., the difference between the measured and modeled DSCDs) arises from an imperfect minimum of the cost function." It is not clear to me what that means. It is hard to imagine that the minimization algorithm (e.g, Gauss-Newton or Levenberg-Marquard-Algorithm) is not working correctly and does not yield a minimum of the cost function. Do you suspect that the algorithm ends up in a local minimum of the cost function that is different form the global cost function? This can happen, but you would need to show that this is the case which is difficult to do. Usually, one distinguishes between the following error components [Rodgers, 2000], which should be discussed accordingly:

1. Smoothing error caused by the limited vertical resolution due to the limited information content of the measurements. These are discussed in the manuscript based on the averaging kernels.
2. Forward model errors due to imperfect representation of the physics of the system. In case of profile retrieval, this could for example be caused by horizontal inhomogeneities of trace gases and aerosols.
3. Retrieval noise caused by the noise of the measurements
4. Forward model parameter errors, caused by errors or incomplete knowledge of model parameters, such as the aerosol phase function.

Re: Thank you for this comment. As you said, it might be inaccurate to describe like "…arises from an imperfect minimum of the cost function". Algorithmic error points to the difference between the measured ($\boldsymbol{y}$) and modeled DSCDs ($F(\boldsymbol{x},\boldsymbol{b})$) as follows.

$$\sigma_{algorithm} = \mathbf{y} - F(\mathbf{x}, \mathbf{b})$$

where $F(\boldsymbol{x}, \boldsymbol{b})$ is the forward model; $\boldsymbol{b}$ represents the meteorological parameters; $\boldsymbol{y}$ is the measured DSCDs; $\boldsymbol{x}$ is the state vector.

In an ideal situation, the modeled DSCDs should equal the measured ones. The error sources that result in this difference include forward model error from an imperfect approximation of forward function F, forward model parameter error from selection of parameter $\boldsymbol{b}$, and errors not related to the forward function parameters (e.g., detector noise). In other words, algorithmic error contains forward model error and forward model parameter errors, which are the main contributors, but not identical to the sum of them. On the other hand, the forward model error is hard to be quantified due to the difficulty of acquiring an improved forward model.

We have changed our demonstrations in Section 2.4 and Supplementary Sect. S3 as follows.

Section 2.4: "Algorithm error (i.e., the difference between the measured and modeled DSCDs) mainly arises from an imperfect representation of the real radiation field in the RTM - spatial inhomogeneities of absorbers and aerosols, clouds, real aerosol phase functions etc."

Supplementary Sect. S3: "Algorithm error is the discrepancy between the measured ($\boldsymbol{y}$) and modelled DSCDs ($F(\boldsymbol{x}, \boldsymbol{b})$). As displayed in Eq. s3, the error sources that result in this discrepancy include forward model error from an imperfect approximation of forward function F, forward model parameter error from selection of parameters $\boldsymbol{b}$, and errors not related to the forward function parameters, like detector noise (Rodgers, 2004). Algorithm error is a function of the viewing angle. Due to the difficulty of assigning this error to each altitude of profile, the algorithm errors on the near-surface values and column densities are usually estimated by calculating the average relative differences between the measured and modeled DSCDs at the minimum and maximum elevation angle (except 90°), respectively (Wagner et al., 2004).

$$\sigma_{algorithm} = \mathbf{y} - F(\mathbf{x}, \mathbf{b}) \quad \text{(s3)}$$

where $F(\boldsymbol{x}, \boldsymbol{b})$ is the forward model; $\boldsymbol{b}$ represents the meteorological parameters; $\boldsymbol{y}$ is the measured DSCDs; $\boldsymbol{x}$ is the state vector."

References

Chance, K. and Spurr, R. J. D.: Ring effect studies; Rayleigh scattering, including molecular parameters for rotational Raman scattering and the Fraunhofer spectrum, Appl. Opt., 36, 5224–5230, https://doi.org/10.1364/AO.36.005224, 1997.

Grainger, J. F. and Ring, J.: Anomalous Fraunhofer line profiles, Nature, 193, 762, https://doi.org/10.1038/193762a0, 1962.

Rodgers, C. D.: Inverse methods for atmospheric sounding, theory and practice, edited by: Taylor, F. W., World Scientific, 2000.

Rodgers, C. D., Taylor, F. W. (Ed.): Inverse methods for atmospheric sounding, theory and practice, World Scientific, 255 pp.2004.