

Reply to Ref. #1

First of all we want to thank this reviewer for the positive assessment of our manuscript and the constructive and helpful suggestions.

General comments

Yang Wang et al. presented a comprehensive study of the temporal and spatial distribution of aerosols and trace gases in the central-western North China Plain. The manuscript is well structured and the results show good correlation with results of other instruments.

However, many results were filtered based on a cloud classification scheme which performs somehow unreliable to me. Especially low and high aerosol loads seem to be mixed up with clouds. To a certain points, MAX-DOAS profiles should be able to retrieve these differences but do not show larger deviations e.g. on 11/05 and 19/05. Since this classification has a large impact on the complete discussion, I would suggest to add a small estimate of the impact of wrongly classified/filtered scenarios on your results.

Furthermore, an additional analysis of NO₂ retrieved in a different fitting window might help to clarify if horizontal inhomogeneties were present.

Author reply:

Many thanks for the positive assessment! We modified the manuscript based on the comments from you and the other two reviewers. The one-to-one replies are given in the following.

For your comments on cloud classification results, first of all, we agree on that some cases with clouds in the reality might be identified as “clear sky with high aerosols” or “low aerosols”, due to certain thresholds are used. The problems can occur if the quantities which are used for the cloud identifications are close to these thresholds. Following your suggestion, we discussed the issue in the Sect. 2.2.4 of the revised manuscript as following:

“Since certain thresholds are used for the identification of cloud scenarios, two sky conditions might be interchanged because the derived quantities are close to the chosen thresholds. The problem occurs relatively often between the ‘cloud free and high aerosol load’ and ‘continuous clouds’ categories because they are only distinguished by the absolute value of the color index. The issue can impact the MAX-DOAS results of aerosol profiles and AODs due to the remaining cloud contamination. Fortunately the problem can be easily solved if an additional filter is applied, which is the convergence between measured and modelled O₄ dSCDs in the profile inversion for aerosols, based on the previous study in Wagner et al. (2016). If the convergence is bad, the corresponding aerosol results are possibly contaminated by clouds. Therefore the filter of convergence is applied to the MAX-DOAS results for the statistical analysis and elaborated in section 2.2.5 and Table 2.

In addition the issue can also impact the comparisons of MAX-DOAS results with coincident independent measurements under different sky conditions in section 4.1. However since the cases close to the thresholds do not dominate in each category, the general conclusions on the effects of clouds and aerosols are not significantly impacted.”

Meanwhile we also modified the other part of Sect. 2.2.4 to illustrate the cloud classification scheme more clearly for the readers.

However for the mixing-up of the different sky condition on 11/05 and 19/05 which you pointed out, we think the phenomenon is not certainly due to the wrongly classified scenarios, but probably due to the true temporal variations of the sky conditions. You also pointed out that the

MAX-DOAS (aerosol) profiles do not show large deviations between the different sky conditions. This finding actually indicates the necessity of the cloud classification. The sky classification scheme is mainly based on the color index, which is ratio of intensities at 330nm against those at 390nm, while aerosol profile retrievals are based on O₄ absorptions. Our previous study indicates that usually the color index is more sensitive to sky conditions than the O₄ absorptions. This explains why the sky change of the condition changing is hardly seen from the aerosol profile results but is clearly seen from cloud classification results. Therefore the cloud-contaminated aerosol results need to be filtered out based on the cloud classification results.

Regarding horizontal inhomogeneous distributions of NO₂ and retrievals of NO₂ in the visible range, please see my reply to the specific comment 26.

Specific Comments:

1) P2, L14-19: Please add a reference to Fig 1. in the Introduction.

Author reply: we modified the manuscript by adding a reference to Fig 1 regarding the NCP region.

2) P2, L32: Please add the full name for the abbreviation East-Aire.

Author reply: The full name is East Asian Study of Tropospheric Aerosols: an International Experiment and added in the manuscript.

3) P3, L9: MAX-DOAS algorithms are not only based on OE. Iterative approaches like Newton-Gauß or Levenberg-Marquardt are in use. I would rather call these algorithms "inversion algorithms" or "inversion based algorithms".

Author reply: We modified the manuscript based on the suggestion as follows: "... inversion algorithms based on the optimal estimation (OE) method (and iterative approaches, e.g. the Newton-Gauß or Levenberg-Marquardt, are also used)"

4) P5, L17: Wang et al 2018 seems to be the wrong reference?

Author reply: Sorry for the mistake. The wrong reference is deleted in the revised manuscript.

5) P5, L26: "and for the MAD-CAT campagin" → "e.g. the MAD-CAT campaign"

Author reply: we modified the revised manuscript accordingly.

6) P5, L30: You state detection limits but not how they were calculated. Please add the missing information.

Author reply: Thanks for the suggestion! The detection limits are estimated based on the typical DOAS fit errors of individual species. The information is added in the revised manuscript.

7) Section 2.2.3: The uncertainties of the individual parameters were given as percentages but where do they come from and how were they calculated? From a previous study or from not shown sensitivity tests? Please give some information.

Author reply: Thanks for pointing this out! The uncertainties of the profile retrievals due to a potential bias of the SSA and the Ångström exponent are derived based on sensitivity tests, which are not shown in the manuscript. We modified the manuscript to illustrate the information as the following:

“A systematic bias of the SSA typically contributes to an uncertainty of about 5% to the retrieved aerosol and trace gas profiles from MAX-DOAS measurements. These values were derived from sensitivity tests by varying SSA in the profile inversion.”

“The uncertainty of the Ångström exponent (due to uncertainties of sun-photometer measurements) typically contributes to uncertainties of up to 20% to the retrievals of trace gas profiles. These results are derived from sensitivity tests by varying the Ångström exponent between 0.49 to 2.53 in the profile inversion. For the assumed range of the Ångström exponent see the discussion in section 3.1.1. ”

- 8) P6, L11-12: When there is a sunphotometer measuring routinely. Why not using the exact SSA and asym parameter closest in time rather than averaged values? For which wavelengths are the averaged quantities? How did you convert to the proper wavelengths or did you assumed no wavelength dependency?

Author reply: The reason why we used the averaged values of SSA and the asymmetry parameter is due to the measurement uncertainties of the sun-photometer. Both parameters depend on the aerosol type, which is often similar for larger periods. Also there are many gaps in the measurement time series of the SSA and the asymmetry parameter due to the cloud filtering and quality controlling. In addition, both parameters are not measured in the UV spectral range, but retrieved at 440nm from the sun-photometer measurements. We add the missing information in the revised manuscript as follows:

“A fixed single scattering albedo (SSA) of 0.95 and an aerosol phase function parameterised according to Henyey and Greenstein (1941) with an asymmetry parameter of 0.72 are chosen according to averaged inversion results at 440nm from the sun-photometer also operated at the measurement station.”

“It needs to be clarified that considering uncertainties of inversions of the SSA and asymmetry parameters of sun-photometer measurements, average values of both parameters are used in the inversion of MAX-DOAS measurements.”

- 9) P6, L17: How were these wavelengths chosen? 354 for HONO is the mid of the fitting window but what about the other wavelengths? Was there are reason for not choosing the fitting window mid wavelengths?

Author reply: The wavelengths are the effective wavelengths of air mass factors of individual species in individual spectral ranges of the DOAS fits. The effective wavelengths can be calculated by weighting the wavelengths by the differential cross section values as shown in the previous study of Marquard et al. (2000) (Marquard, L. C., Wagner, T., & Platt, U. (2000). Improved air mass factor concepts for scattered radiation differential optical absorption spectroscopy of atmospheric species. *Journal of Geophysical Research*, 105(D1), 1315–1327. <https://doi.org/10.1029/1999JD900340>). We clarified this point in the revised manuscript as follows: “The air mass factors simulated by RTM are used for the profile inversion and the simulation wavelengths are calculated by weighting the wavelengths by the differential absorption cross section within the individual spectral ranges of the DOAS fits based on the method elaborated in the previous studies, e.g. Marquard et al. (2000).”

- 10) P6, L21: Why was the upper grid limit chosen to be at 3km? Typical altitudes from other studies are usually at 4km.

Author reply: Thanks for the asking! This is a mistake. In order to clarify the point, we modified the manuscript in the beginning of section 2.2.3 as the following:

“Tropospheric vertical profiles of aerosol extinction and volume mixing ratios (VMRs) of NO₂, SO₂, HONO, HCHO, and CHOCHO are retrieved from the elevation-dependent dSCDs by using the PriAM profile inversion algorithm (Wang et al., 2013a, b, 2017a) with a vertical grid of 200 m in an altitude range of up to 4 km. From the derived profiles the vertical column densities (VCD) of the trace gases and AODs are derived by vertical integrations. Due to the fact that no substantial information on the concentrations above 3km can be derived from the measurements, the retrieved profiles below 3 km are shown in all figures of the study.”

11) P6, L25: Covariances of 100% of the surface value for all altitudes? Is this correct? The commonly used approach is a fixed percentage of the a priori profile of the individual altitudes.

Author reply: Thanks for pointing out this mistake! The manuscript is modified as:

“The diagonal elements of the a-priori covariances (S_a) at different altitudes are set as the square of 100% of the a-priori values at individual altitudes in order to balance the flexibility and stability of the profile inversion.”

12) Table 2: Why is there a different SZA limit for SO₂ compared to the other trace gases?

Author reply: The reason is that SO₂ is retrieved at a shorter wavelength range compared to the other species. The intensity decreases stronger along increasing of SZA and the interference with the O₃ absorption in the DOAS fit of SO₂ is much stronger at a high SZA. We clarified this point in the revised manuscript as the following:

“Here it needs to be noted that a lower SZA threshold is set for the filtering of the SO₂ results than for the other species, because the intensity at short wavelengths is rather low and spectral interferences with the O₃ absorption increases strongly with SZA.”

13) P6, L30: Here, you write R but in the Figure its R². Which one was given?

Author reply: Sorry for the mistake. The correlation parameter shown in the figures is R². In order to be consistent with other figures, we modified the figures and show R instead.

14) Fig 4: Please change the colors for either low or high aerosols in this and similar plots because it is hard to distinguish between both markers.

1. The aerosol retrieval shows similar profile shapes from 6 to 13 but the cloud classification finds different cloudy conditions, sometimes with thick clouds. How is it possible that the aerosol retrieval is not affected by thick clouds?

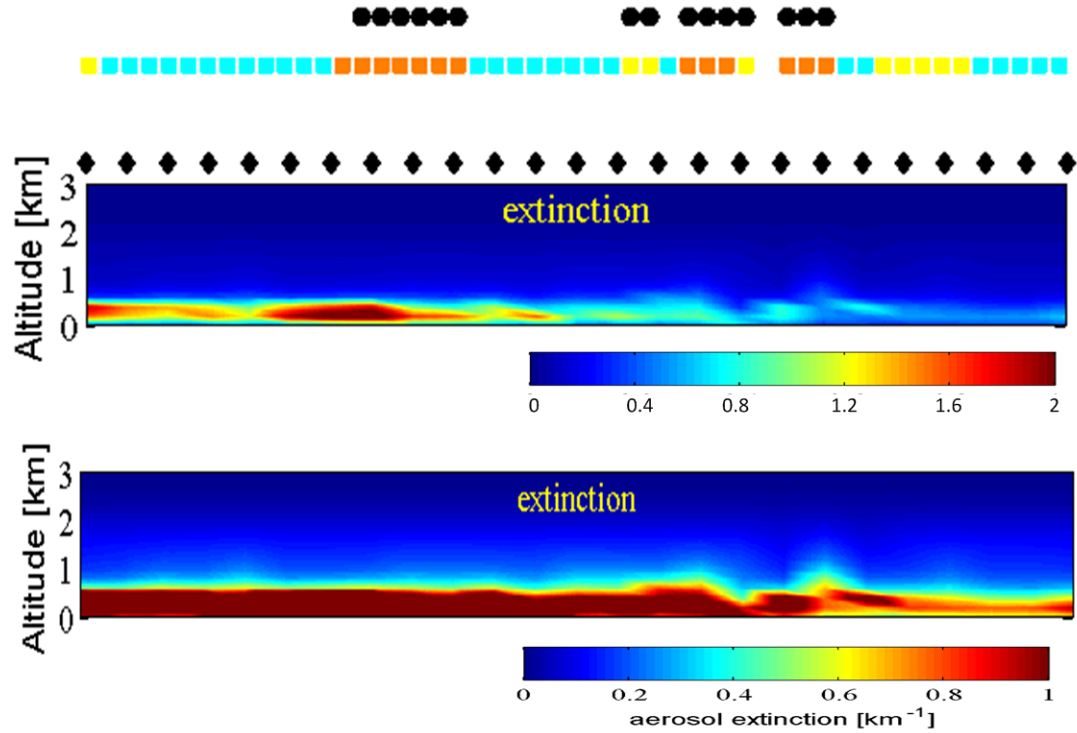
2. The a priori profiles for aerosols and SO₂ are not even close to the retrieved profiles. How can you be sure that you do not over- or underestimate the retrieved profiles due to an inaccurate a priori profile?

3. I do not understand why the degrees of freedom for aerosols are larger than any of the trace gas ds. This is unexpected for me. Could you please explain where these larger differences in ds for the individual retrievals come from?

Author reply: the “cloud free with low aerosols” conditions do not appear on that day. Therefore we think it should not be a big problem to distinguish the blue and light blue colors on this day.

For your first comment, the maximum value of the color bar of 1 km⁻¹ is partly misleading. If the maximum value is enlarged to 2 km⁻¹, the effects of clouds can be clearly seen. The updated

figure is shown below. In addition, large extinctions at high altitudes can not be retrieved with certainty from MAX-DOAS measurements under optically thick and continuous clouds, because similar light paths can be expected at different elevation angles. In such cases, because the differential O₄ dSCDs at off-zenith views compared to the zenith view are used in the inversion of aerosol profiles, the differential O₄ dSCDs coming from the contributions of clouds are close to zero. Therefore no information on clouds can be derived from the differential O₄ dSCDs in such cases. However the identification of optically thick clouds is mainly based on the O₄ dSCDs in the zenith view compared to those under cloud-free sky conditions. Therefore the thick cloud conditions can be identified from MAX-DOAS.



For your second comment, we compared the AOD and near-surface values of SO₂ with the other co-located near-surface measurements. The comparisons verified the MAX-DOAS retrievals. We also show the comparisons of simulated and measured dSCDs in the supplement. If the profile inversion can well reproduce the dependences of dSCDs on elevation angle, the retrieved profiles are probably close to the truth. Further if MAX-DOAS measurements provide sufficient information on profiles, the retrieved profiles are expected to be different from the a-priori. The dependence of the profile inversion algorithm should on a-priori profile should be small. In the PriAM, we do the inversion in the logarithmic space and using the non-linear Levenberg-Marquardt iterative approach in order to reduce the constraints of a-priori.

For your third comment, since ds is the sum of the diagonal elements of the averaging kernel (A), and A can be calculated as the following equation:

$$A = \frac{\partial x}{\partial x_i} = GK = (S_a^{-1} + K^T S_\epsilon^{-1} K)^{-1} K^T S_\epsilon^{-1} K$$

For aerosol retrievals, K is the response of O₄ dSCD to a variation of the aerosol extinction. Therefore A is a function of the aerosol extinction. For trace gas retrievals, K is the response of the trace gas dSCDs to the logarithmic concentration of the trace gas. Therefore A is a function of

the actual aerosols and the trace gas profiles, and A can be expected to be different for individual measurements and species. The reason why a larger d_s is found for the aerosol retrieval than for the trace gases for the noon measurement shown in Fig. 4 is probably the fact that aerosols extend to higher altitudes than the trace gases. Actually a large d_s can be also seen for the CHOCHO inversion (CHOCHO also extends to high altitudes).

15) P7, L15: Why is the near-surface extinction trusted under partially cloudy conditions but the AOD not? I would also assumed that the near-surface extinction is inaccurate when broken clouds led to a contamination of some elevation angles only. The profile will be smoothed due to the a priori smoothing and the retrieval for all altitudes should be affected.

Author reply: The filtering scheme is based on our previous study of long-term comparisons of MAX-DOAS aerosol results with sunphotometer and visibility meters in Wang et al. (2017a) (Wang, Y., Lampel, J., Xie, P., Beirle, S., Li, A., Wu, D., and Wagner, T.: Ground-based MAX-DOAS observations of tropospheric aerosols, NO₂, SO₂ and HCHO in Wuxi, China, from 2011 to 2014, *Atmos. Chem. Phys.*, 17, 2189-2215, <https://doi.org/10.5194/acp-17-2189-2017>, 2017a.). The relevant figure in the paper is given below. It indicates that the systematic differences of the MAX-DOAS results of the near-surface aerosol extinction compared to the visibility-meter under all cloudy sky conditions are similar to those under cloud free sky conditions. We agree with you that broken clouds can impact near-surface aerosol results for individual measurements. Since the effects are different if clouds are observed at different elevation angles, and the impacted elevation angles are random, their contributions to obviously cancel out for the averages of long term measurements. In order to elaborate the point more clearly, we added the following information in the revised manuscript:

“Here it needs to be noted that clouds, especially broken clouds, can impact the MAX-DOAS results of the near-surface aerosol extinction for individual measurements. However since the cloud effects occur for different elevation angles, their overall impact is rather random. Therefore if long term measurements are averaged, cloud effects on the near-surface aerosol mostly cancel out and do not contribute to a systematic bias (Wang et al., 2017a).”

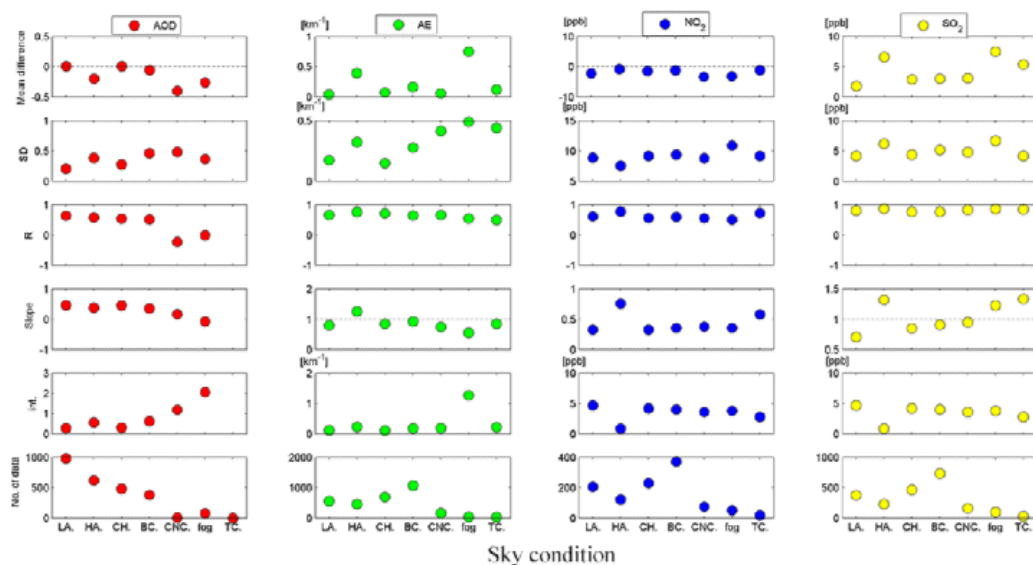


Figure 9. Mean absolute differences and standard deviations as well as correlation coefficients (R), slopes and intercepts derived from linear regressions of the AODs, and near-surface AE, NO_2 and SO_2 VMRs between MAX-DOAS and independent techniques for different sky conditions in autumn. The corresponding number of data points is shown in the bottom panel. Different colours denote AOD (compared with the Taihu AERONET level 1.5 data sets), AE (compared with the nearby visibility meter) and NO_2 and SO_2 (compared with the nearby long-path DOAS instrument). For calculations of mean absolute differences, the data derived from independent techniques are subtracted from those derived from the MAX-DOAS instrument. The data derived from the MAX-DOAS instrument are plotted against those derived from independent techniques for linear regressions.

16) P7, L25-28: The numbers are averages over the full time period? Please call it a total average then or specify what these numbers exactly mean.

Author reply: The numbers are averaged over the whole campaign period. Therefore we modified the sentence in the revised manuscript as suggested.

17) P8, L3-5: 5-times the word "also" in three lines. Maybe you can reformulate these two sentences.

Author reply: The sentences are modified in the revised manuscript.

18) P10, L6-7: I am confused of what you wrote in Section 2.2.3 and what you wrote in lines 6-7 (see comment P6, L11-12). Please explain your approach in greater detail.

Author reply: We modified the sentence in the revised manuscript as follows: "The Ångström parameter, single scattering albedo, and asymmetry factor are also retrieved from the sun-photometer measurements and are used as input for the inversion of the aerosol profiles from MAX-DOAS measurements, for details see section 2.2.3. "

In addition, we modified section 2.2.3 in the revised manuscript based on your comment P6, L11-12.

19) P10, L10: A parameter of 1 means an Angström Exponent of 1? Please replace the word "parameter". Furthermore, was the exponent of 1 used for all data or was the sunphotometer's Angtröm exponent closest in time applied for the conversion?

Author reply: We changed “parameter” to “Exponent”. As mentioned above (your comment 8), considering the retrieval uncertainty of the sun-photometer measurements, the averaged value during the whole campaign period is used. We modified the sentence in the revised manuscript as follows:

“Aerosol extinction at 360 nm is derived from the visibility at 550 nm using an Ångström exponent of 1, which is the average value derived from the sun photometer measurements during the whole campaign period.”

20) P11, L28-29: In addition to your reason, a general lower sensitivity for higher altitudes might be another reason. Furthermore, the limitation to 3km might also be an issue.

Author reply: We agree on that a general lower sensitivity for higher altitudes can also play a role here. Therefore we added the following sentence to the revised manuscript:

“Low sensitivity of MAX-DOAS measurements on aerosols located at high altitudes can also play a role. This finding is illustrated in Fig. 8b and discussed in section 4.3.”

Concerning the effect of the limitation to 3km (see our reply to your comment 10), this was a wrong information. The inversion is done in the altitude range below 4km.

21) P12, L1-2: Generally, I would not assume that an elevated layer within the lowest kilometre can not be resolved by a MAX-DOAS profiling algorithm. For higher altitudes, this might be an issue. Is it possible to add a brief synthetic test on 3 or 4 elevated layers in different altitudes? Just to see the retrieval response. Because an elevated layer could also be possible when the underlying aerosol profile is box-like due to an oscillation around this box-features.

Author reply: We agree the explanation of the findings in the manuscript was not convincing. Based on this comment (and the comment 22), we modified the explanation as follows:

“The largest underestimation of the MAX-DOAS results compared to the visibility-meter results is found in the morning and when aerosol loads are large. Since the boundary layer height is lower in the morning, larger vertical gradients of aerosols in the layer between 0 and 200m can be expected. The different air mass measured by the MAX-DOAS and the visibility-meter and a large vertical gradient near the surface might be the reason for the large underestimation of the MAX-DOAS results compared to the visibility-meter results.”

22) Fig 6:

1. It would be interesting to know if these outliers for high visibility-meter AE correspond to certain geometries, time or weather conditions?

2. P12, L11: When this behaviour for different cloudy conditions can be attributed to clouds, why is there a different correlation for NO₂ and SO₂ for high aerosols and cloudy sky? NO₂ has a better correlation for cloudy sky while SO₂ has a better correlation for high aerosols.

Do you expect large inaccuracies in the classification scheme for clouds and aerosols?

Author reply: For your comment 1, many thanks for the good suggestion to see the dependence on the time of the day! We modified Fig. 6 in the revised manuscript by adding colors to show the time of the day. For the comparison of the near-surface AEs, we can see that the largest underestimation of MAX-DOAS results compared to visibility meter results occurs in the morning and when the aerosol load is large. Since the boundary layer height is lower in the morning, larger vertical gradient of aerosols can be expected. In the lowest vertical layer between 0 and 200 m, since aerosols might accumulate near the surface. The visibility meter measures aerosols near the surface, but the MAX-DOAS inversion results represent averages of the

aerosols extinction in each layer. The different air masses measured by MAX-DOAS and the visibility-meter together with large vertical gradients in the lowest layer between 0 and 200 m might be the reason for the underestimation of the MAX-DOAS results compared to the visibility-meter results. The revised manuscript is modified as the following:

“The comparison of the near-surface aerosol extinction between the MAX-DOAS measurements and the visibility-meter is shown in Fig 6b. While very good agreement is found for the category “clear sky with low aerosols” (R of 0.81 and slope of 1.02), much worse correlation is found for “clear sky with high aerosols” (R of 0.32 and slope of 0.14). Even worse agreement is found for the category “cloudy sky” (R of 0.22 and slope of 0.11). The largest underestimation of the MAX-DOAS results compared to the visibility-meter results is found in the morning and when aerosol loads are large. Since the boundary layer height is lower in the morning, larger vertical gradients of aerosols in the layer between 0 and 200m can be expected. The different air mass measured by the MAX-DOAS and the visibility-meter together with strong vertical gradients near the surface might be the reason for the large underestimation of the MAX-DOAS results compared to the visibility-meter results. In addition the effect of clouds on the MAX-DOAS aerosol retrievals are probably the reason for the larger scattering under the cloudy sky conditions compared to cloud-free sky conditions. Note that clouds are not included in the RTM which is used for profile retrievals of aerosols and trace gases.”

For your comment 2, the different cloud effects for NO₂ and SO₂ can be attributed to their different vertical distributions. As shown in Fig. 12, SO₂ typically extends to higher altitudes than NO₂. Therefore cloud effects, especially for broken clouds, on the near-surface SO₂ concentration can be expected to be stronger than for NO₂. Moreover, for NO₂, the photolysis rate can be expected to be lower under cloudy conditions than under cloud-free conditions. Therefore the vertical distributions of NO₂ in the layer between 0 and 200 m might be smoother under cloudy conditions than under cloud-free conditions. The smoother distribution might be the reason for the better agreement of the NO₂ results between MAX-DOAS and in-situ measurements. Although clouds can also affect the NO₂ profile inversion, since NO₂ is close to the surface, the effect of a smoother vertical distribution of NO₂ under cloudy conditions might be important for the comparison of MAX-DOAS and in-situ results of NO₂. In order to clarify this point, we added the following text in the revised manuscript:

“Further, different cloud effects on NO₂ and SO₂ comparisons are found. For SO₂ worse correlation is found under “cloudy sky” conditions than under “clear sky with high aerosols”, while a better correlation is found for NO₂. Effects of clouds in the profile retrievals of SO₂ can be expected to be stronger than for NO₂ due to the fact that SO₂ can extend to higher altitudes than NO₂ as shown in Fig. 12. In addition, the photolysis rate of NO₂ can be expected to be lower under cloudy conditions than under cloud-free sky conditions. Therefore the vertical distributions of NO₂ in the layer between 0 and 200 m might be smoother under cloudy conditions than under cloud-free sky condition. The smoother vertical distribution is probably the reason for the better agreement of the NO₂ results between the MAX-DOAS results and in-situ measurements.”

23) P12, L12-14: A limited vertical sensitivity for near-surface trace gas concentrations might also be important when the trace gas is concentrated in a shallow layer much smaller than the grid step width of the profiling algorithm, especially when using a coarse grid steps width of 200m.

Author reply: We agree on the suggestion. We want to say the same thing with the sentence of “This finding is probably related to vertical and horizontal inhomogeneity of the species”. In

order to elaborate it more clearly, we added a further explanation in the bracket as “(especially in the lowest vertical layer between 0 and 200 m)” in the revised manuscript.

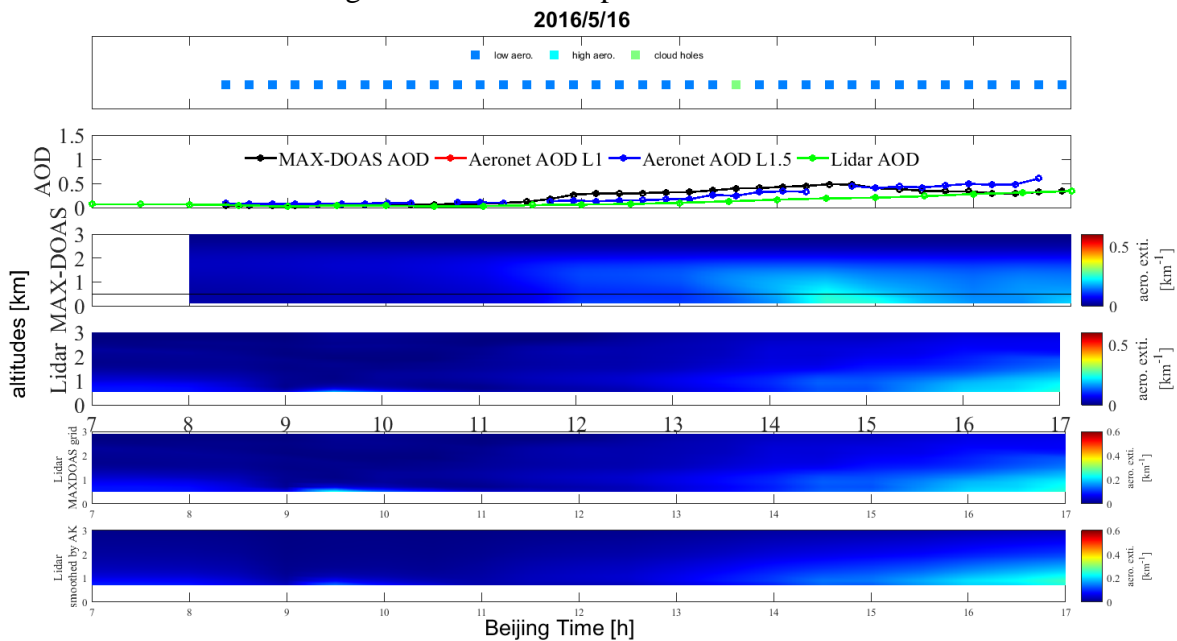
24) Section 4.2:

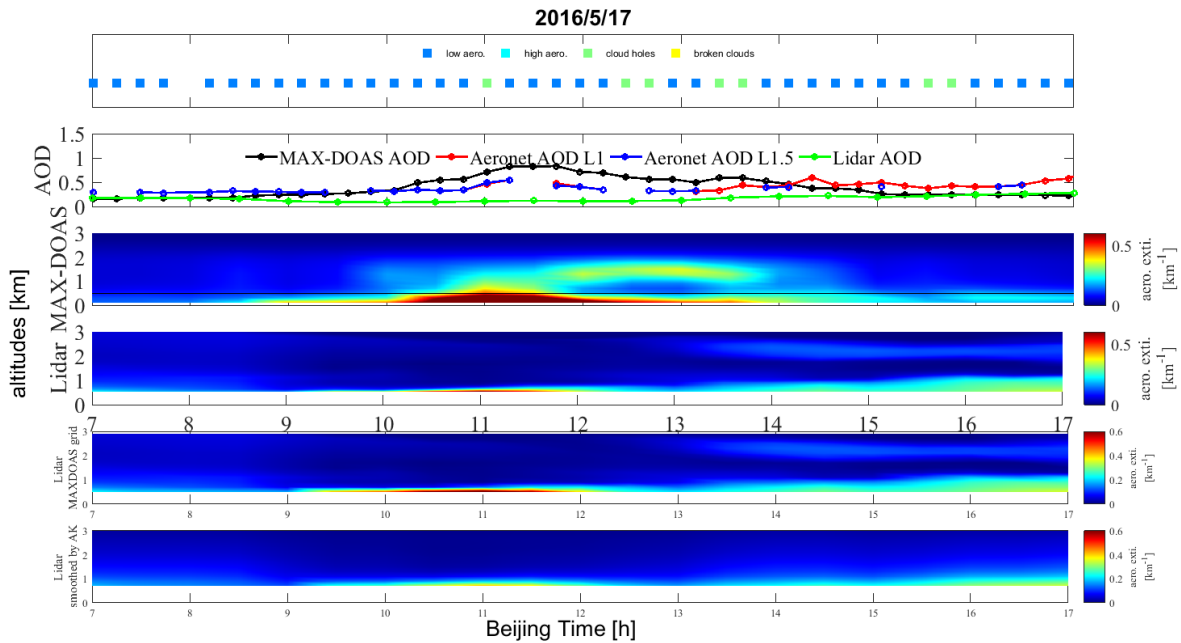
1. What is the vertical resolution of the Lidar?
2. Was AK smoothing applied for the Lidar measurements (with the assumption that the vertical resolution is higher than that for MAX-DOAS)? Please add also the Lidar AOD to the AOD sub-figure.
3. Why are high near-surface values not similarly found by both instruments? The sensitivity for near-surface values should be the highest for both instruments and different air masses and clouds are not expected to be that important for lower altitudes. E.g. MAX-DOAS found larger extinctions at 15:00 (2016/5/16) while the Lidar found larger values in the evening of 2016/5/17.

Author reply: 1) The vertical resolution of the Lidar measurements is 7.5m. The information is added in the revised manuscript.

2) The Lidar results shown in Fig. 7 are the original results without the smoothing by the AK of the MAX-DOAS measurements. Since you asked for it, we plot the Lidar profile smoothed by an AK in the bottom subfigure of the plots below. We also show the lidar profiles interpolated to the grid of MAX-DOAS profile retrievals in the fifth subfigure of the plots below. Since the plots of smoothed profiles do not give more important information compared to the original profiles in the comparisons with MAX-DOAS results, we decide not to show the smoothed profiles.

The AODs (green) calculated from the original Lidar aerosol profiles are also given in the plots below. Since there is a blind area of the lidar measurements below 500m, the AODs derived from the Lidar profiles are not suitable for the quantitative comparison with the MAX-DOAS AOD results due to large amount of aerosols located at altitudes below 500m. Therefore we prefer not to show the Lidar AODs in Fig. 7 of the manuscript.





3) First of all, different air masses are measured by the MAX-DOAS and Lidar instruments. MAX-DOAS pointed to the North with a typical horizontal effective light path of about 5 to 10 km. In contrast, the Lidar pointed to the zenith. Horizontal inhomogeneous distributions of aerosols might play a role on the differences between the two techniques. Secondly, as you mentioned in your previous comment, broken clouds can impact some elevation angles of MAX-DOAS measurements. The same broken clouds are probably not be seen by the Lidar in the zenith. The effects of inhomogeneous cloud coverage can also contribute to the differences between the two techniques. The potential reasons for the differences have been given in the manuscript as follows:

“The remaining differences can probably be explained by the fact that different air masses are observed by both techniques, while horizontal inhomogeneities of aerosols and cloud cover could appear. For example, different clouds and aerosols could be observed by both instruments. Note that the MAX-DOAS telescope was pointed towards the North, while the Lidar measured the atmosphere directly above the station. The sun-photometer measured the air masses in the direction of the sun.”

25) P13, L8: How did you calculate these "combined profiles"? Linear interpolation between lowest air-craft and surface value? Please add some information.

Author reply: Your understanding is correct. A linear interpolation is used. We added the information in the revised manuscript as follows:

“Combined profiles” are generated by combining averaged aircraft profiles with averaged surface measurements. Values at altitudes between the lowest aircraft measurements and surface measurements are generated by linear interpolations.”

26) Fig 8:

1. Please add times for MAX-DOAS measurements and the overpasses. It would also be nice to have some color-coding or different grey-scales for the green surface values to identify the surface value changes throughout the measurement period.

2. Why do the SO₂ curves always agree better than NO₂ even though the degrees of freedom are much lower for SO₂? Did you try to retrieve NO₂ also in another fitting window (> 400nm)? It would be interesting to see if the results differ strongly. That would support the argument of horizontal inhomogeneities.

Author reply: 1) The times of the comparisons are added in the revised manuscript and figure. Regarding color-coding of the in-situ data, we prefer not to follow this suggestion, because the current plots are already quite busy. The motivation to show the in-situ data over the comparison period is to illustrate the temporal variability of the pollutants near the surface. Thus the time information might be not important here. In addition if the readers want to see the time series of the in-situ measurements, they can find them in Fig. S3 of the supplement.

2) An important reason why the aircraft NO₂ results are much larger than the MAX-DOAS results might be the interference of NO_y with NO_x in the aircraft measurements. The same effect can be seen from the comparison with the surface in-situ measurements of NO₂ shown in Fig. 6c. The same technology is applied to the aircraft and ground based measurements. We added this information in the revised manuscript as follows:

“For NO₂, significantly larger values of the aircraft measurements than the MAX-DOAS measurements can be seen on 21 May. Since the same technique as for the ground based in-situ measurements of NO₂ (see section 4.1) was used in the aircraft measurements, the interference of NO_y with NO_x might also cause an overestimation of the NO₂ concentrations derived from the aircraft measurements.”

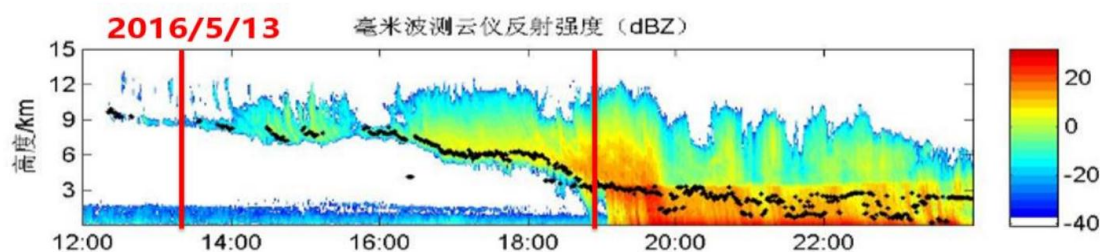
We tried to retrieve NO₂ in the visible range, the general time series are consistent between the UV and visible range. I think it needs a sophisticated inversion algorithm to retrieve horizontal inhomogeneity from NO₂ absorptions in the UV and Visible. This is beyond the scope of the study.

27) P13, L24-25: Please add the time when the photos were taken on the individual days.

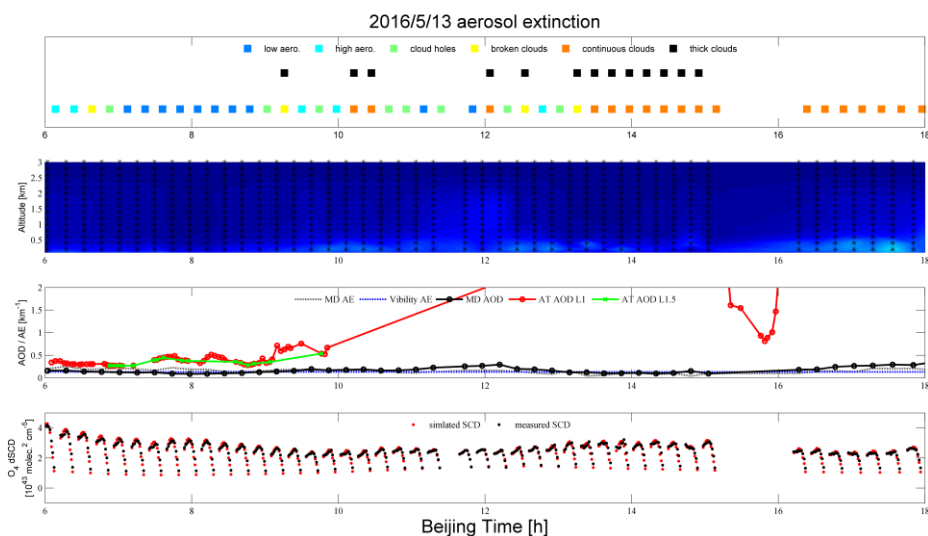
Author reply: the photos were taken around noon on all the days. The information is given in Fig. 9.

28) Section 5.1: Why does the cloud classification show highly variable results between 6 and 14 BT (11/05) but the aerosol profiles do not differ strongly? This indicates that either the profiles or the classification is inaccurate. In addition, around noon two days later (13/05), thick clouds were found but the aerosol retrieval does not show these clouds. This is surprising.

Author reply: For the question regarding 11/05, the answer has been given in the reply to comment 14. The same reason can explain the findings on 13 May. If the continuous clouds are located at high altitudes, similar absorption paths of O₄ in the cloud for the zenith view and off-zenith view of MAX-DOAS measurements can be expected. Therefore with the zenith spectrum as the FRS in the O₄ DOAS fits of off-zenith spectra, the partial O₄ dSCDs due to the contribution of the cloud will cancel out. Therefore no clouds can be retrieved under such conditions using the O₄ dSCDs at off-zenith views. We show the backscattering signal measured by a collocated ceilometers in the figure below. It indicates that the bottom height of clouds on 13 May is around 8km. By comparing the cloud classification results with the ceilometers results, we can conclude that the cloud classification results are correct.



We present the comparisons of measured and modelled O₄ dSCDs derived from aerosol profile retrievals of MAX-DOAS in the bottom of the figure below. We can see good agreement of the measured and modelled O₄ dSCD. Therefore we conclude that the results of the profile inversion should be realistic.



29) P15, L21-22: Trajectories ending at different altitudes were used, but how? You do not give information about that. Did you just average all maps for Figure 11?

Author reply: Your understanding is correct. We generated the maps with trajectories ending at different altitudes. Then the maps were averaged to generate Fig. 11. In order to clarify this procedure, we added the following text in the revised manuscript:

“In order to consider pollution transport at different altitudes, we generated individual maps using trajectories ending at 100m, 300m, 500m, 1km, and 2km above the measurement site. These maps are then averaged to generate the final map. The individual maps using trajectories ending at different altitudes are presented in Fig. S4 in the supplement. The final maps are shown in Fig. 11.”

30) Section 5.2.1: It would be interesting to see how these maps change when the individual lifetimes are considered. Are CTM calculations of lifetimes at these days available?

Author reply: We showed the maps with different backward times in Fig. S4 of the supplement. We use different backward times for the generation of the maps to implicitly test the effects of varying lifetime. Note that no information about the lifetime is available from the CTM in this study. In addition, we don't think that the method could be in general be improved if information on the lifetime would be used to scale the trace gas VCDs for the generations of the maps, because the trace gas columns measured by MAX-DOAS are actually transported from areas in

different distances. It is unknown which fraction of the total trace gas amount was transported from an area in a specific distance. Therefore the scaling of the VCD based on the lifetime can even cause artificially high VCDs in the generated maps if large backward time is used.

31) Fig 12: Please add the data of the validating instruments, when possible.

Author reply: Here we did not follow the reviewer's suggestion, because we feel that the plots would be too busy, while at the same time no relevant information will be added. Here it should be noted that the comparisons with independent measurements have been already shown in Fig. 6. Moreover, the study of the effects of transports is only based on the MAX-DOAS results.

32) P17, L14-15: How can I differ between the original data and the interpolated data in Fig 13?
Please use other marker styles.

Author reply: Sorry for the missing information in the manuscript. The data shown in Fig. 13 are all original data. No interpolated data are shown. Therefore we delete the sentence "Note that the data are interpolated to provide a more consistent overview. The original data are also plotted in Fig. 13 in order to show the representativeness of the interpolated data." in the revised manuscript.

33) P17, L16: Highest values for southeast winds with southerly trajectories instead of north-westerly?

Author reply: The higher values are found for southeast winds when the long-range trajectories are from the northwest. This finding indicates the effect of short-range transport of pollutants emitted in the downtown area of Xingtai (with an iron factory) about 20km southeast of the station. The effect of short-range transport can only be seen when the air mass over the station is dominated by transport of clean air masses from the northwest. In order to describe it more clearly, we modified the sentence in the revised manuscript as follows:

"In general the higher values of aerosols, NO₂, SO₂, and HCHO occur for the southeast wind directions than for other wind directions on the days of north-westerly trajectories. This finding possibly indicates the effect of short-range transports of pollutants emitted in the downtown area of Xingtai (with an iron factory) about 20km southeast of the station. The effect of short-range transport can be well identified when the air mass over the station is dominated by transport of clean air mass from the northwest."

34) P18, L18-19: of on the order of hours --> in the order of hours

Author reply: It is corrected in the revised manuscript.

35) Fig 14c: The red dots are extremely small. Please increase the size of these dots.

Author reply: The images with the red dots are downloaded from <https://worldview.earthdata.nasa.gov>. Therefore we can't change the size. We added the website for Fig. 14c in the revised manuscript.

36) P31: Wang, F et al.: Please change 2048 to the proper year.

Author reply: It is modified as "2018" in the revised manuscript.