Response to comments by Reviewer #1

We thank Reviewer #1 for the detailed review of our manuscript. The suggestions and comments helped us to further improve the revised version. Below we copy each comment and respond to each of them in blue fonts.

This manuscript evaluated the relationship between total-column and near-surface NO2 in winter in Beijing in using in-situ and PANDORA ground-based remote sensing observations. Detailed vertical distribution was obtained from Lidar observations. Then the possible influence of the atmospheric boundary layer evolution on the near-surface pollutants was discussed. Also, the coupling mechanism between the transport of pollutants and the complex vertical distribution of the atmospheric boundary layer was analyzed using back trajectories and weather maps. The manuscript is well-written and fits the scope of the journal. It provides an interesting case study of the vertical distribution changes of pollutants in the boundary layer based on remote sensing technology. However, there are a few aspects that need to be addressed to further improve the quality of the manuscript.

1. The manuscript highlighted the use of PANDORA for stratification inversion. This is an interesting point, but a detailed explanation of its credibility as well as its role in practical applications are lacking. Please further elaborate on the advantages of PANDORA technology to state its importance and accuracy to readers.

Response: Thank you for your comment. Here, based on the technical methodology outlined in the ATBD, we briefly introduce the process of the gas profiling retrieval algorithm using PANDORA as follows:

The first step in determining the partial column amounts of NO_2 in different layers is to estimate the effective height (center) corresponding to a given Pointing Zenith Angle (PZA_i), which is based on the air-gas slant columns using geometrical relationship and assuming single scattering conditions only:

$$h_{CENi} = \frac{\Delta qs_{i,AIR} + q_{CLIM,AIR}}{n_{SURF,AIR}} \cdot \frac{cos(PZA_i)}{2}$$

Where, $\Delta qs_{i,AIR}$ is the absolute slant column amount of the air-gas at PZA_i. $q_{CLIM,AIR}$ is absolute vertical column amount of the air-gas from a climatology. $n_{SURF,\,AIR}$ is the climatological surface concentration of the air-gas. The measured differential slant columns at n_{PZA} angles (i=1 to n_{PZA}) give us n_{PZA-1} atmospheric layers. The top height of layer i is given by:

$$h_{TOPi} = 0.5 \cdot (h_{CENi+1} + h_{CENi})$$

The lowest layer extends from the surface to h_{TOP1}, where the corresponding angles are PZA_{MAX} and PZA_{MAXm1}. The next step is to calculate differential air mass factors for NO_2 and the air-gas for each layer, Δm_i and $\Delta m_{i,AIR}$ respectively:

$$\Delta m_i = \frac{qs_i - qs_{i+1}}{q_T}$$

$$\Delta m_{i,AIR} = \frac{qs_{i,AIR} - qs_{i+1,AIR}}{q_{CLIM,AIR}}$$

Where, qs_i and $qs_{i,AIR}$ is the absolute slant column amount of the NO2 and air-gas at PZA_i. The profile shape p_s of the partial columns is determined as a variation of the air-gas shape:

$$ps_i = \frac{0.5 + \Delta m_i - \Delta m_{i,AIR}}{1 + \Delta m_{i,AIR}} \cdot q_T$$

Where, q_T is the tropospheric column. For the normalized profile shape psn, the profile shape is divided by its largest element, hence all psn_i are $\leqslant 1$. The average number density of the NO_2 in layer i, n_i, is then calculated from the following empirical equation:

$$n_i = \frac{2 \cdot (\Delta qs_i - \Delta qs_{i+1}) + 0.5 \cdot q_T \cdot (m_{i,RAYL} - 3 \cdot \Delta m_{i,AIR})}{\Delta qs_{i,AIR} - \Delta qs_{i+1,AIR} + q_{CLIM,AIR}} \cdot n_{SURF,AIR} \cdot psn_i$$

$m_{i,RAYL}$ is the Rayleigh air mass factor for the corresponding PZA. It is calculated from Radiative Transfer Calculations. Further correction due to profile differences are addressed by the correction factor cf_i:

$$cf_i = psn_i \cdot \frac{n_i}{Max(n_i)}$$

Max() refers to the maximum over all i. For those layers with negative values of cf_i the n_i are corrected in the following way:

$$n_i = n_i \cdot (1 + cf_i)$$

Negative values of n_i are set to 0. For the partial column amounts Δq_i, the concentrations n_i are multiplied with the layer thickness:

$$\Delta q_i = n_i \cdot (h_{TOPi} - h_{TOPi-1})$$

For i=1, the lowest layer, h_{TOPi-1} equals 0. The last step is to normalize the Δq_i with the tropospheric column:

$$\Delta q_i = \Delta q_i \cdot \frac{Sum(\Delta q_i)}{q_T}$$

Sum() refers to the sum over all i.

Regarding the accuracy of the NO_2 TranProf product at four levels (0-4 km), both official PANDORA documents, "Fiducial Reference Measurements for Air Quality" and "Pandonia Global Network Data Products Readme Document," clarify that there is currently no accuracy assessment for this product. When using it, we followed the accuracy screening criteria for L1 products and retained the sections with high accuracy in the radiometric observations.

The NO_2 profile inversion is explained by modifying the text at the end of Sect. 2.2.1 to:

Line 225-241: “Diffuse (scattered) radiation is measured at 5 pointing zenith angles (PZAs) in sky

mode which, together with the direct sun measurement, provides information on the tropospheric VCD and on the surface concentrations. The PZAs are 0º, 60º, 75º, 88º and a maximum angle taken as 89º. The measurements are taken in a V shape (all angles are measured twice around a central angle) as described in Cede (2024). Four partial columns of NO_2 concentrations are provided by the PANDORA inversion. The first step is the estimation of the effective height corresponding to a given PZA, and then calculate differential air mass factors for the NO_2 and the air-gas for each layer. The profile shape of the partial columns is determined as a variation of the air-gas shape. The average number density of the NO_2 in each layer is then calculated. The partial column amounts can be obtained from the concentrations multiplied with the layer width as described in the Manual for Blick Software Suite (Cede, 2024), Section 6.7. The NO_2 of the partial column can be obtained from the uvh3 product which was downloaded from the PGN website (https://pandonia-global-network.org, last accessed: 22 Jan 2025). We converted these partial column concentrations into layer-averaged volume mixing ratios and interpolated them to 6 standard levels (0.2, 0.5, 1.0, 1.5, 2.0, 2.5 km) for visualization."

2. The authors used total concentrations when analyzing NO2 obtained from ground-based remote sensing, but compared the tropospheric concentrations of NO2 with that from satellite sensor TROPOMI. This needs to be clarified.

Response: Thank you for your comment. Firstly, we used TROPOMI satellite observations of tropospheric NO_2 to confirm that the pollutants originated from regional pollution transport processes, rather than to strictly compare satellite data with ground-based measurements. Secondly, we reviewed the TROPOMI satellite products and found that only tropospheric products were available during the time period of our study. Finally, in our previous research (Liu et al, 2024), we have compared total vertical column densities (VCDs) and tropospheric VCDs of NO_2 from TROPOMI with PANDORA using 2022 data, and obtained good validation results. For clearer description in the manuscript, we have added an explanation as follows:

Line 269-273: "In addition, only tropospheric vertical column densities (VCDs) for NO_2 were available during the time period of this study. The TROPOMI tropospheric VCDs for NO_2 are only used as a qualitative reference for upwind concentrations for the evaluation of effects of long-range transport using air mass backward trajectories and not for quantitative analysis."

3. The observation was divided into three periods, in which PM2.5 continued to rise during Period II and suddenly dropped dramatically on the 25th. In the case study, the authors only chose the two cases of Period I. The cases during Period II were not analyzed. Please explain this in detail.

Response: Thank you for your comment. There was indeed a persistent heavy pollution process during Period II. However, due to the occurrence of clouds during that period (as mentioned in Sect. 3.1 , line 340 "total VCD of NO_2 observations are not available due to the presence of clouds", we are unable to obtain available observations in cloudy or severely hazy weather conditions. This is a limitation of the use of optical remote sensing and we were unable to select case studies from Period II. We have stated this reason in the manuscript as follows:

Line 413-418: "Our study primarily focuses on the correlation between total VCD and NS concentrations of NO_2. Total VCD of NO_2 are provided by the passive remote sensing instrument PANDORA, which does not provide reliable observations on cloudy days, as mentioned above. This is the reason why we only selected cases during Period I and did not obtain cases during Period II,

when pollution was more severe."

4. In the case studies of the 14th and 18th, the PANDORA remote sensing only had observations during the daytime. This makes sense. However, there was a difference in the vertical distribution of NO2 from PANDORA versus Lidar. I can understand that the Lidar signal comes more from the scattering of aerosol particulate matter, but the vertical decoupling of NO2 does not appear in the PANDORA observations. Although the authors have some explanation for this, I think it needs further clarification of these differences.

Response: Thank you for your constructive comment. PANDORA retrieval provides NO_2 column concentrations by dividing the atmospheric column into 4 layers, for which the reported height represents the top of each layer. In Figures 3d and 5d of the manuscript, we converted these column concentrations into layer-averaged volume mixing ratios (from mg m^{-2} to mg m^{-3} units) and interpolated them to 6 standard levels (0.2, 0.5, 1.0, 1.5, 2.0, 2.5 km) for visualization (for more detail see our response to your 1st comment). To clarify the vertical resolution limitations of PANDORA, we now present the original 4-layer data in Figure R1. The results show: When NO_2 column concentrations are elevated, the retrieved layer top heights decrease significantly, with high-value concentrations primarily distributed between 0.2-1.0 km and increasing with time during a day. Lidar observations reveal complex sub-kilometer stratification within the 1 km boundary layer, with a resolution of 7.5 m, where PANDORA's coarse 4-layer resolution limits its ability to resolve fine-scale NO_2 gradients. This was mentioned in the original manuscript on lines 777-778: “with differences attributed to the vertical resolution of the Pandora and lidar observations”. For further clarification, we have added the following explanation to the manuscript:

Lines 512-517: " These differences stem from the difference in vertical resolution between the PANDORA and the lidar: the total VCD from PANDORA is divided into two layers (approximately 300-800 m and 800-1700 m) within the detailed stratified height range (300-1000 m) observed by lidar. Consequently, the fine stratified structure within 1000 m cannot be identified with the available PANDORA data."

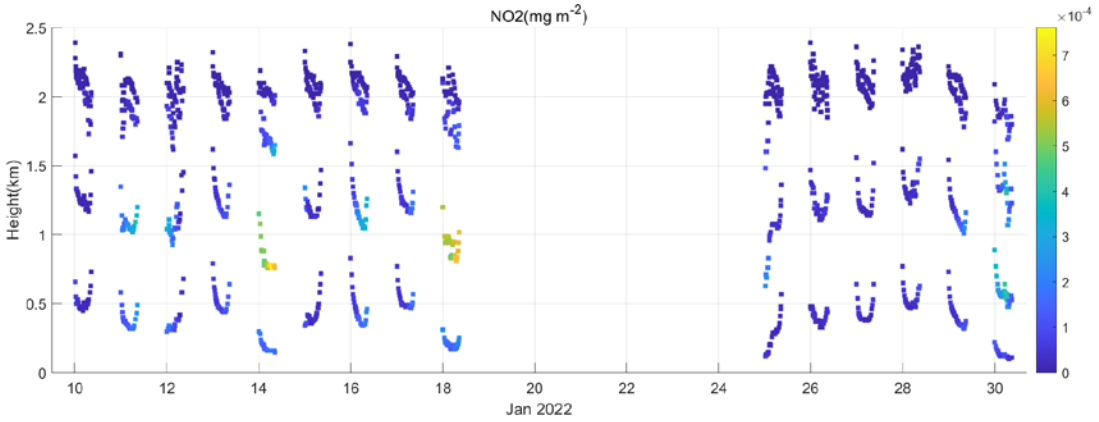

Figure R1. NO_2 column concentrations retrieved by PANDORA

5. Based on the multi-temporal backward trajectory and weather maps and in conjunction with the remote sensing imagery from TROPOMI, both cases were in good agreement with the regional transport pattern. In the vertical distribution evolution, it was also consistent with the expectation of winter boundary layer development. Although the authors have explained the boundary layer

development and the conformity of the backward trajectories at different heights, the differences between the two cases were not yet clearly summarized. Despite the presence of multiple layers in the boundary layer in both mornings, there is a correlation between the distribution of pollution concentrations in the layers and the regional pollution intensity and transport direction, which is presented in both cases. It would be great if the authors could verify whether this is the case and summarize the differences, which I believe would be interesting.

Response: Thank you for your comment. We did not clearly articulate the basis for selecting the cases or their characteristics. In fact, we chose typical cases based on the pollutant transport paths in Beijing. Following Yin et al. (2025), SW (Shanxi-Beijing) and SE (Shandong-Hebei-Beijing) are two common pollutant transport pathways affecting Beijing.

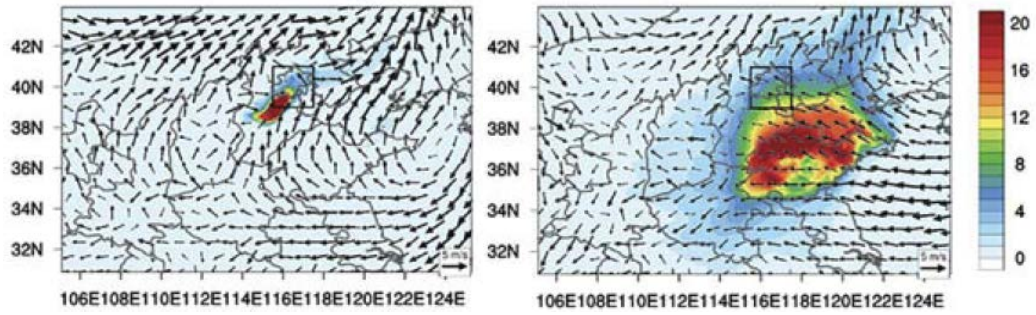

Fig R1. The wind field and pollutant absolute contribution from different regions under SW and SE patterns. (Yin et al., 2025)

Case 1 (January 14th) represents a belt-shaped pollution event mainly transported from Shanxi to Beijing, as indicated by the elongated SE-NW pattern (Fig. 4a). The 24-hour air mass trajectories (Fig. 4c) confirm the overall transport from the SW, with the lower two trajectories close to the surface during the first 12 hours. These two lower trajectories (arriving at 300 and 500 m), clearly indicate the stagnating air mass, but coming from different direction during the last few 6-hour steps: the air mass trajectory arriving at 300 m travelled from the NE during the last 12 hours while the air mass arriving at 500 m travelled even shorter distances during the last 18 hours and arrived from the NW. These trajectories clearly show the separation of air masses at different heights, passing over different areas and thus with different influences as regards picking up pollutants from different sources. And possibly these pollutants are influenced by different processes since wind speeds were different, and other influencing meteorological factors may have been different as well.

Case 2 (January 18th) features a large-scale pollution event extending from the southern part of Beijing into Shandong, Hebei and Henan) (Fig. 6a). The air mass trajectories (Fig. 6c) show the overall transport from the WNW over large distances (10°) for the air mass arriving at 1000 m, but with changing patterns for the airmasses arriving at 300 and 500 m. Initially, at 10:00, these air mass show the stagnating air, with very low windspeed from the S (300 m) to the SW (500 m). At 13:00, the 500 m air mass trajectory originated from 1000 m (as opposed to the air mass arriving at 10:00 that originated near the surface and travelled at 500 m) and had travelled about 8° from the west during 18 hours when it stagnated during the last 6 hours near the arrival location. The air mass arriving at 300 m had travelled a very short distance from the SSW and stagnated near the arrival point. The air masses arriving at 16:00 all had travelled longer distances (8° or more) and arrived from the W during the last 6 hours and earlier from the W (300 m), WNW (500 m) and NW (1000 m). This clearly marks the end of the pollution episode with increasing wind speeds, transporting cleaner air from the W and NW and ventilating the pollution out. Also in this case the air mass

trajectories indicate very different origins of the air mass at different heights, and thus different pollution sources.

In Case 1, higher pollutant concentrations were observed in the 800-1000 m layer (Fig 3a) because the air mass primarily transported from Shanxi to Beijing along the SW direction, while the pollutants were carried by the air mass as it passed through Shijiazhuang City in Hebei Province. In Case 2, pollutants were mainly transported from the plains, resulting in higher pollutant concentrations in the near-surface layer. These characteristics align well with the spatial distribution and vertical stratification of pollution, and both cases are highly representative. We have summarized the characteristics of the two cases as follows:

Line 672-699: "A comparative analysis of the two cases reveals distinct characteristics. Case 1, on January 14th, is characterized by a belt-shaped pollution event. The air mass arriving at 1000 m was primarily transported from Shanxi to Beijing arriving from the SW during the last 15 hours (Fig. 4c), while at lower levels the air masses travelled at low altitudes (from near the surface to 500 m) through the polluted area in Hebei (compare Figs. 4a and 4c). In contrast, Case 2, on January 18th, was a large-scale pollution event covering Shandong, Hebei, and Henan and the southern part of Beijing (Fig. 6a), but the air mass arriving at 1000 m had been transported form the WNW over a large distance over clean areas. However, at lower levels, the air was stagnant (wind speed was low) in the morning, air masses arriving at 300 and 500 m at 10:00 LT had travelled over very short distances during the last 24 hours and were thus only influenced by local pollution. In Case 1, elevated pollutant concentrations were recorded in the 800-1000 m altitude layer in the morning, attributable to an air mass originating from Shanxi Province (with no significant pollution observed there) that was transported over the polluted area in Hebei at about 1000 m (Fig 4a and 4c), and then the upper-level air mass likely carried pollutant residuals which had been uplifted through vertical mixing processes over the polluted area in Hebei. These pollutants were subsequently advected to Beijing, where their presence was detected by lidar observations. Conversely, in Case 2, pollutants were predominantly transported from the plains, leading to a significant accumulation of pollutants in the near-surface layer. After 13:00, in both cases the distinction between the pollutant layers disappeared when the boundary layers developed under the influence of surface heating and increasing wind speed (Fig. 2), thus creating boundary layer turbulence and mixing of NO_2, aerosols and other constituents. Both the total VCD and NS concentrations of NO_2 increased, with that of total VCD of NO_2 being more significant. This further suggests that it is more difficult to obtain NS NO_2 concentrations using total VCD of NO_2 concentrations during the morning hours. However, utilizing the types of pollution spatial distribution and transport patterns can be helpful in indicating NS NO_2 concentrations."

6. I have some confusion about Figure 7. Although it showed regular data, the ratios were not the correlation coefficients or slopes that were used in the case studies. The slopes were fixed in the cases where TC correlates well with NS NO2 concentrations, whereas the ratios were not slopes, which were highly variable. I think the authors need to give a more detailed explanation of how they are related and different.

Response: Thanks for your comment. We apologize for this confusion and try to explain better why we introduced Ratio. The analysis of the two cases in Section 3.2 shows the good correlation between total VCD and NS NO_2 concentrations in the afternoon, and for Case 1 also in the morning, but not for Case 2. This prompted us to look also at other days to determine whether and when a good correlation occurs and whether a linear relationship between total VCD and NS NO_2

concentrations, characterized by y = ax + b, can be established. This exercise showed that for a larger number of days the correlation is less good than in cases 1 and 2, particularly during the morning hours. The low correlation does not quantify the changes between total VCD and NS NO_2, prompting us to introduce the Ratio as a means of quantification. To further explore a relation between total VCD and NS concentrations of NO_2, we introduce the Ratio between these variables, Ratio = total VCD / NS, and looked at the variation of Ratio. The data shows that the variability of Ratio is generally smaller in the afternoon than in the morning, which indicates the feasibility of using polar-orbiting satellites with an afternoon overpass time to determine NS NO_2 based on total VCD of NO2. However, using geostationary satellites for the same prediction based on total VCD observations before 13:00 poses greater challenges. Notably, the analysis of cases 1 and 2 showed that regional pollution transport paths may be helpful for predicting NS concentrations of NO_2 in winter in Beijing. In other words, considering both the spatial distribution of pollutants and their transport path ways in predictions has the potential to enhance the ability of satellites to predict NS concentrations NO_2 based on observations of total VCD of NO_2. The limitation of our study is that it is based on only one month of observations, and during this period, the geostationary satellite GEO-KOMPSAT-2B /GEMS did not release any products, which prevented us from better demonstrating the universality and applicability of our findings. In future research, we will obtain more observations and utilize geostationary satellite observation products to further refine and validate our discoveries.

We further elaborated on the role of Ratio in the MS as follows:

Line 706-712: “The Ratio, defined as the ratio of total to NS NO_2 concentrations, serves a dual purpose: it not only quantifies the changes between total VCD and NS concentrations of NO_2 when the correlation is low but also reflects the degree of dispersion between the two measurements. A more variable Ratio indicates higher dispersion and poorer correlation, providing a straightforward yet effective way to assess the reliability of using total VCD of NO_2 to predict NS NO_2 concentrations.”

We further summarized the comprehensive point presented by the analysis of Ratio and the two cases as follows:

Line 737-749: “Generally, the temporal stability of the Ratio is important. The Ratio is overall less variable after 13:00, suggesting that polar-orbiting satellites can be used to predict NS NO_2 based on total VCD of NO_2 during this period with greater confidence. This temporal stability is particularly valuable because it offers a feasible approach for air quality monitoring and forecasting. In contrast, the Ratio is less stable before 13:00, posing greater challenges for using geostationary satellites for the same prediction task. It's worth noting that our analysis in winter in Beijing suggests that considering both the spatial distribution of pollutants and their transport direction has the potential to enhance the ability of satellites to predict NS NO_2 concentrations based on total VCD of NO_2. By incorporating this information into prediction models, the accuracy and reliability of satellite-based air quality predictions may be improved, particularly in complex urban environments where pollutant concentrations can vary significantly over short distances and time periods.”

7. There are also some other minor issues that need further improvement, such as the format of the citation.

Line 167: “… north, east, and west (Figure 1). ”, for figures, the manuscript used fig more often, please make sure that they are consistent.

Response: We have modified it to "… north and west (Fig. 1). " and checked the whole manuscript.

Line 205: "… precision of 0.01 DU and a nominal accuracy of 0.1 DU Herman et al. (2009)", the citation should be in parentheses as a whole.

Response: The manuscript has been revised.

Line 260: "PM2.5 is measured using beta rays generated", "2.5" should be subscripted.

Response: The manuscript has been revised.

Line 304: "e.g., Atkinson, 2000; Boersma et al., 2009; Y. Zhang et al., 2016", "Y. Zhang" should be "Zhang". Please double check all the citations so that the format is consistent.

Response: The manuscript has been revised.

Line 442: "… with a wind speed of 2 m/s", for units, superscripts are used more often than slashes throughout the manuscript. Please revise it to be consistent.

Response: The manuscript has been revised.