Reply to referee #1

The authors thank the referee for the valuable time spent to thoroughly read the manuscript and provide valuable comments which contributed to improvement of this revised version. Below we provide our point-to-point responses, together with the revisions made, where appropriate.

(Referees' comments in red, author responses in black, and adjustments of manuscript in blue.)

The authors report the NO_2 observations in Beijing China from first Pandora instrument, show the local temporal variation of NO_2 and reveal the spatial and temporal representativeness of atmospheric column NO_2 concentrations obtained from ground-based remote sensing. The manuscript is well structured and logic, gives some observational facts and valuable conclusions, and deserves to be published in the AMT journal. However, I still have a few comments as below I hope authors will clarify before publication.

1. In section 2.2, two subheadings 2.2.3 appear. Also the reanalysis data are not instrumental and should not be presented in this section; the authors are requested to adjust them.

   Thank you for this comment, the second 2.2.3 was revised to 2.2.4, and we changed the section heading: '2.2 Instrumentation' to '2.2 Instrumentation and auxiliary data'.

2. This manuscript focuses on the analyses and comparisons of pandora instrumental observations and is not solely a measure of the differences between TROPOMI and pandora observations; The paper shows that the pandora observations are also compared to ground-based observations at least and that the differences are measured. It is therefore recommended that the methods section be revised and improved by correcting the description of the paragraph below line 225 and adding the description of the methods in the other sections, if any.

   Thank you for this comment. We fully agree that we use both satellite data and in situ observations in this study. This is also clearly mentioned throughout the manuscript (abstract, introduction, and all other Sections 2, 3 and 4, with subsections to Sections 2 and 3 devoted the parts of the study where satellite and/or in situ data are used). All this is included in "Evaluation" as we used it in

the title and we specifically mention “satellite validation” in the title as an additional aspect because this is an important reason for establishing the Pandora Network. For the evaluation by comparison with independent data sets, like satellite and in situ data, it is common to use statistical metrics and these are summarized in Section 2.3, after a discussion of collocation criteria. Realizing that the correlation coefficient had been mentioned but not defined, we have added a brief description to the text in line 241:

“

The Pearson correlation R (Pearson., 1895) is defined in Eq. (1).

$$R = \frac{\sum_{i=1}^{n}(VCD_{TROPOMI,i}-\overline{VCD_{TROPOMI}})(VCD_{Pan,i}-\overline{VCD_{Pan}})}{\sqrt{\sum_{i=1}^{n}(VCD_{TROPOMI,i}-\overline{VCD_{TROPOMI}})^2}\sqrt{\sum_{i=1}^{n}(VCD_{Pan,i}-\overline{VCD_{Pan}})^2}}, \quad (4)$$

”

The description of the collocation and statistical metrics used is applied in several sections and therefore should be in methods, to avoid repeating.

According to the Referee#1’s comments, we have added some words at the beginning of Section 2.3 which now reads “For the evaluation of the NO_2 observations different methods are used, such as time series to show the variability on different time scales or the effects of external parameters such as wind speed, averaging to reduce short-time variability, scatterplots for comparison with independent data sets. For the comparison between Pandora and TROPOMI NO_2 VCDs, the data need to be collocated.”

3. In line 252 for the ratio of DQ2 data to total data, 2176 divided by 80,153 does not equal 28.2%. Please check and revise.

Thanks for this comment. We have rechecked the numbers of data in line 261, and 21767 was misspelled as 2176. The associated text in manuscript has been revised to " Among the total VCDs, 21767 data points out of a total of 80153 (27.2%) are low quality. ".

4. In Figure 2 we can see that the number of observations changes from month to month. How is this variation taken into account in the statistical process, e.g., by calculating the median, mean, etc.? Do you average all the observations within a time period in a month or divide the data into per days first and then take the mean?

Thank you for this comment. In our manuscript, indeed we first divided the data into per day and then take the mean to avoid effects on the monthly means due to differences in the number of days when observations were available. This is now also clarified in the text by adding “monthly mean tropospheric VCDs

(calculated from daily averages)" (line 294 and in the caption of Figure 3 we have added "Monthly mean data were calculated from daily means.".

5. In section 3.2, the authors may have missed a phenomenon. There are still several red dots distributed in the north-west around the interval 270° to 320° in Fig 4. However, the author states that clean air is transmitted from the northwest. I think this may not be a coincidence and would appreciate an explanation.

Thank you for this comment. Indeed a small number of red dots, indicating high NO_2 concentrations, occur in the North-west wind sector. We have looked at these data points in more detail and came up with the following explanation which has been included in the text (lines 340-347) "As a result, the NO_2 concentrations in the northwesterly wind sector are generally low, as shown in Figure 4. However, the data in Figure 4 show some exceptions when NO_2 concentrations are high. Further analysis shows that these observations were all made during the winter and are likely due to NO_x emissions from natural gas companies located in the Changping district in the northwest of Beijing. Natural gas is provided for, e.g., heating in the winter, and NO_x is produced during the combustion process (Pan et al., 2023). Thus, in the winter, during northwesterly winds, NO_2 is transported to the Pandora site. This explains the observations of high NO_2 concentrations, due to local emissions. (More details of high concentration number and time during north-west wind please see Table S2.)"

Table S2: Period of high NO_2 concentration during north-westerly winds during winter.

Time	Tropospheric NO_2 VCD (Pmolec.cm^{-2})
2021/11/14 12:32:22	32.8
2021/11/17 13:16:32	48.7
2021/11/24 12:44:22	31.7
2021/12/14 11:27:36	41.5
2021/12/14 13:09:05	41.5
2021/12/19 13:15:25	33.1
2021/12/22 12:18:21	29.9
2021/12/28 12:05:44	28.6

2022/1/1 12:31:14	25.1
2022/3/8 11:58:46	29.9

6. In section 3.6, why the spatial representation of Pandora is 10km instead of 20km, I noticed that the D_f mentioned in this manuscript is very close between 10km and 20km, with a difference of only 0.002. What is the significance of the author's introduction of D_f if it is not to be used as a metric for evaluation? I would be grateful if this was clarified.

   Thank you for your suggestion. Actually, D_f is a reference metric. In our method, we combine both D_f and standard deviation as an indicator of spatial representativeness. The first step is to do an initial screening of spatial representativeness based on D_f,: D_f did not change between 1 and 10 km (value 1,011 +/- 0.001) but after 20 km D_f had changed to 0.987, i..e the between 10 and 20 km D_f had changed much more (by 0.024), see Table 2. In Table 2 we also see that at 20 km the standard deviation is twice as large as at 10 km, so we consider this not acceptable. In addition, therefore, in Section 3.6, the spatial representation of Pandora is 10km.

Pan, H., Geng, S., Yang, H., Zhang, G., Bian, H., and Liu, Y.: Influence of H2 blending on NOx production in natural gas combustion: Mechanism comparison and reaction routes, International Journal of Hydrogen Energy, 48, 784-797, https://doi.org/10.1016/j.ijhydene.2022.09.251, 2023.