Response to comments on "Maximizing the Scientific Application of Pandora Column Observations of HCHO and NO_2."

We thank the reviewers for comments and suggestions. The manuscript is revised by incorporating their suggestions and comments. We are also thankful to the editors for their time. Comments from reviewers are in black, responses are in blue, and new text added to the manuscript is in bold blue. The page and line numbers referenced correspond to the track changes document.

Editor

Both reviewers recommend accepting your manuscript after minor technical corrections. Please incorporate all their suggestions in your revision.

In particular, Reviewer #2 noted: "The authors have not addressed my concerns about the independence of the data quality assurance methodology." To address this, clearly state that your manuscript presents one approach to data quality assurance not necessarily a universal method and that users may adopt other approaches, as yours is not demonstrated to be the sole QA method.

We now acknowledge that our approach is but one way to filter Pandora observations. We believe it is an improvement on how the community used Pandora observations before the PGN flagging criteria became available. We add the following text to the end of the introduction (Page 3, line 93) and the following paragraph to Page 31, line 608.

The method presented here is but one approach to examine Pandora data quality. The user community should continue to carefully consider how to best ensure data quality while preserving the most data possible.

In response to the overly stringent data flagging procedure, a framework has been described for analyzing Pandora observations by taking advantage of the two independent modes of operation. This includes a new filtering method as an alternative to using quality flags to identify scientifically useful observations. The need to filter data is not new as Pandora users selected filtering criteria with some level of subjectivity in years prior to development of the PGN quality flagging procedure (Herman et al., 2009; Judd et al., 2019; Tzortziou et al., 2022).

RC 1 reviewer comments

General Overview

The presented manuscript proposes an alternative data flagging procedure to the standard PGN flagging for HCHO and NO2 column densities retrieved from MAX-DOAS and direct sun measurements. The aim is to increase the amount of usable data for scientific studies. As such, the topic of the manuscript is important for users of PGN data products. This approach can help data users and readers of the manuscript better understand the standard flagging method and, most importantly, apply their own filter criteria using the presented approach or even go beyond it.

The authors use the linear correlation coefficient as a metric to validate their novel approach for both species, although the primary focus is on HCHO. The correlation of HCHO with surface O3, as well as airborne data for both HCHO and NO2, is presented as a case study.

The manuscript also provides a more in-depth analysis of flagging propagation and identifies the most influential quality indicators responsible for data flagging. One parameter is indeed questionable in terms of whether it should even be used as a flagging criterion. The second parameter is based on spectral fitting RMS. While the authors use an empirical value, the PGN threshold appears to be too strict. However, users should be cautious about disregarding this entirely.

We thank Manuel Gebetsberger for being a referee and providing useful comments and suggestions. To emphasize the importance of the spectral fitting RMS, we add the following sentences to the end of the paragraph on flagging analysis on line 362.

Of these PGN criteria, wrms is the only one considered in the method outlined in Figure 10, but the threshold value is empirically determined independent of the flagging. Data users are encouraged to continue to pay attention to the impact of wrms on Pandora data quality for specific instruments.

Minor comments

Line 160 The 0° direction is not necessarily the ‘preferred’ direction; it is simply a software default in the config file. Therefore, any alignment with an interesting air mass could be coincidental. Or, it may indicate that the instrument owner has not given much thought to the optimal measurement direction. The key message is that, regardless of where the instrument is scanning, the lowest elevation angle should not be obstructed by an obstacle.

We have now revised it as follows on page 7, line 162.

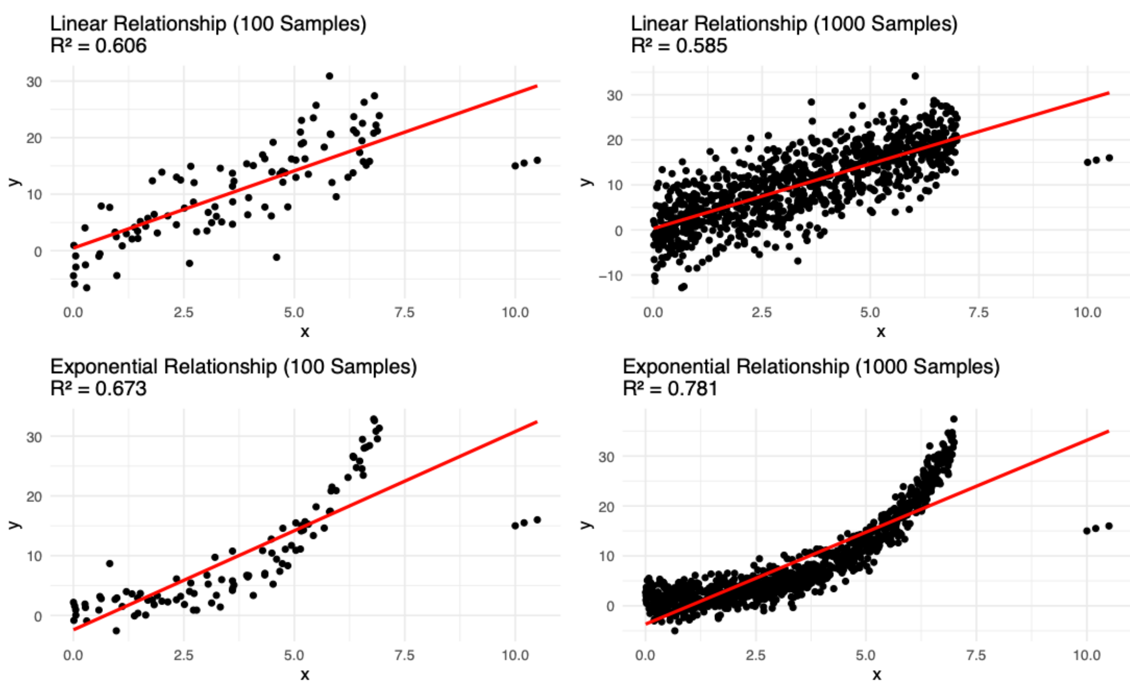
In the SS mode, elevation scans are most effective when taken along an azimuth pointing away from the sun to minimize the effect of scattering.

Line 387 I still see the increase in R2 as being more related to the actual non-linear behavior, which becomes more apparent as the dataset size increases. This can reduce noise in the data and stabilize

R2, which might be overinterpreted compared to the smaller sample (#181) in the SS data, where an outlier has a stronger influence. Since R2 is a linear correlation metric, relying on it alone to assess correlation can be misleading. A larger R2 may simply be an artifact of a larger dataset and does not necessarily imply a stronger linear relationship— especially in cases where the data already exhibit non-linearity and varying variance, as seen here.

The reviewer needs to better explain the perceived nonlinear behavior in Figure 12 which is not evident to us. There is enough scatter in the data to obscure any expectation of specific curvature. As for outliers, there is one obvious one in the SS data for panel (a), but when the amount of data is increased in panel (c), the number of outliers also increases to four points. The influence of removing these outliers is to change the R^2 values in panels (a) and (c) to 0.47 and 0.59 respectively, still indicating a significant improvement with the new filtering method. The main takeaway from these figures is not just the improved R^2, but also increased robustness of the data and utilization of low-quality data, which will enhance the scientific analysis of ozone vs. HCHO column relationships. We have clarified this in our response to the next comment below.

To illustrate this, I have generated a small random dataset with both linear and non-linear relationships, along with some outliers. In the linear sample, the R^2 values remain similar, whereas in the non-linear sample with outliers, the R^2 suggests a 'substantial' improvement.



This example does not appear to be relevant to our dataset. The outliers for the linear data are aligned with the linear trend in the data. For the nonlinear data, the same exact outliers are used, even though it is quite unrealistic to have the same outliers as for the linear case so closely clustered. It is easy to see how this would affect R^2 in the controlled case provided by the reviewer, but there are several things in the data used for this paper that do not follow this example. First, as mentioned above, the HCHO:O_3 data does not exhibit the nonlinear trend in the controlled example. Another is that the very filtering method itself is predicated on an improved R^2 for a smaller (not larger) dataset. If increasing data improves R^2, how is it that our filtering process dramatically improves R^2 when data is removed (see Figures 6 and 7)?

422 Building on the previous comment, in my opinion, the increased R2 does not necessarily improve the analysis of the HCHO:O3 relationship itself. Rather, it appears to be an artifact of the sampling process, combined with the influence of outliers. However, the new filtering increases the dataset size significantly, making other underlying processes more apparent, which may allow for additional conclusions to be drawn.

We do agree that the increase in the size of the dataset is the more important aspect of this work. Accordingly, we have revised the text emphasize the increase in sample size with similar or improved R^2, where we agree that this may make scientific analysis more robust. We revise the paragraph on page 22, line 398 as follows.

Figure 12 shows the relationship between the HCHO column and surface ozone using DS and SS observations. When the new filtering method is applied, **with the number of data points increasing by five-fold for SS (181 to 955) and two-fold for DS (275 to 696)**. **In both cases, the R^2 remains similar or improves. However, the substantial increase in data density will lead to more robust scientific analysis of the HCHO:O_3 relationship**. Similar results are obtained for other months outside the TRACER-AQ period that are not shown here.

Figure S2 For HCHO, there appears to be a slight positive slope, with the difference being around 0 at SZA = 10° (JJA) and approximately 0.4 at SZA = 78° (SON). Such an SZA dependency is not observed for NO2. Is this effect negligible and statistically insignificant for HCHO, or what could be the possible cause of this small slope?

This difference of 0.4E15 is small enough compared to the dynamic range of observed HCHO values such that trying to introduce a SZA-dependent bias correction would not be worthwhile. While we acknowledge the very small slope, we do not have an explanation for it. We do not feel this warrants comment in the manuscript. We have modified Figure S2 for readability to plot the data in decile bins.

Comments on References

Line 677: The provided link is not working. An updated version is available at: Blick Software Suite Manual v1.8.5.

Fixed.

Lines 684–691: The same report is referenced twice. I recommend using the latest ver- sion: PGN Data Products Readme v1.8.10 Please adjust the references accordingly in the manuscript.

Fixed.

Line 716: This reference appears to point to the same link mentioned in my previous comment, likely due to an incorrect copy in the bibliography file.

Fixed.

Line 720: Same issue as the previous two comments. The link appears to be broken and points to the wrong content.

Fixed.

Line 810: Broken link. Please update or remove as needed.

Fixed.

RC 2 reviewer comments

In general, the authors have not addressed my concerns about the independence of the data quality assurance methodology. However, the scientific community have the data and make their own decisions how to use the quality flags and filter the data, so in a sense this paper shows one way to do it and also presents the correlation information between different observations.

See our response to the editor above which addresses this comment.

Here are a few correction recommendations:

Figure 1 (Block diagram of a Pandora instrument and its two viewing geometries) shows the pandora sensor head pointing in the same direction for direct sun and scattered sky. I would recommend changing the angle of the "sky scan" sensor head.

We have changed the angle of the sky scan sensor head in Figure 1.

Figures 8 and 9 have constant Y at changing X while figures 6 and 7 do not. Please explain what the reason is for such difference in NO2 (Fig. 8-9) vs HCHO (Fig 6-7) SS-DS column relationship (if this is not a mistake in selecting DS and SS time coincident data).

In Figures 8 and 9, we display varying X (DS column) values for a constant Y (SS column) for NO_2. This is a result of the contemporaneous matching method, where each SS observations are paired with DS data within a 5-minute window. For NO_2, the frequent DS retrievals enable multiple matches with a single SS observation within this time frame (Figure 2). In contrast, for HCHO measurements, the DS scans are less frequent.

We add the following statement to describe this to page 15, line 310.

Figures 8 and 9 show that for NO_2, the frequent DS retrievals enable multiple DS measurements to be matched with a single SS observation within the 5-minute matching window.