Review_amt-2024-114_v1

Manuel Gebetsberger, LuftBlick

2/18/25

Table of contents

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.

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.

Line 387 I still see the increase in R^2 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 R^2, which might be overinterpreted compared to the smaller sample (#181) in the SS data, where an outlier has a stronger influence. Since R^2 is a linear correlation metric, relying on it alone to assess correlation can be misleading. A larger R^2 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.
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.

422 Building on the previous comment, in my opinion, the increased R^2 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.

```
Warning: package 'ggplot2' was built under R version 4.1.3
```

```
`geom_smooth()` using formula = 'y ~ x'`geom_smooth()` using formula = 'y ~
x'`geom_smooth()` using formula = 'y ~ x'`geom_smooth()` using formula = 'y ~ x'
```

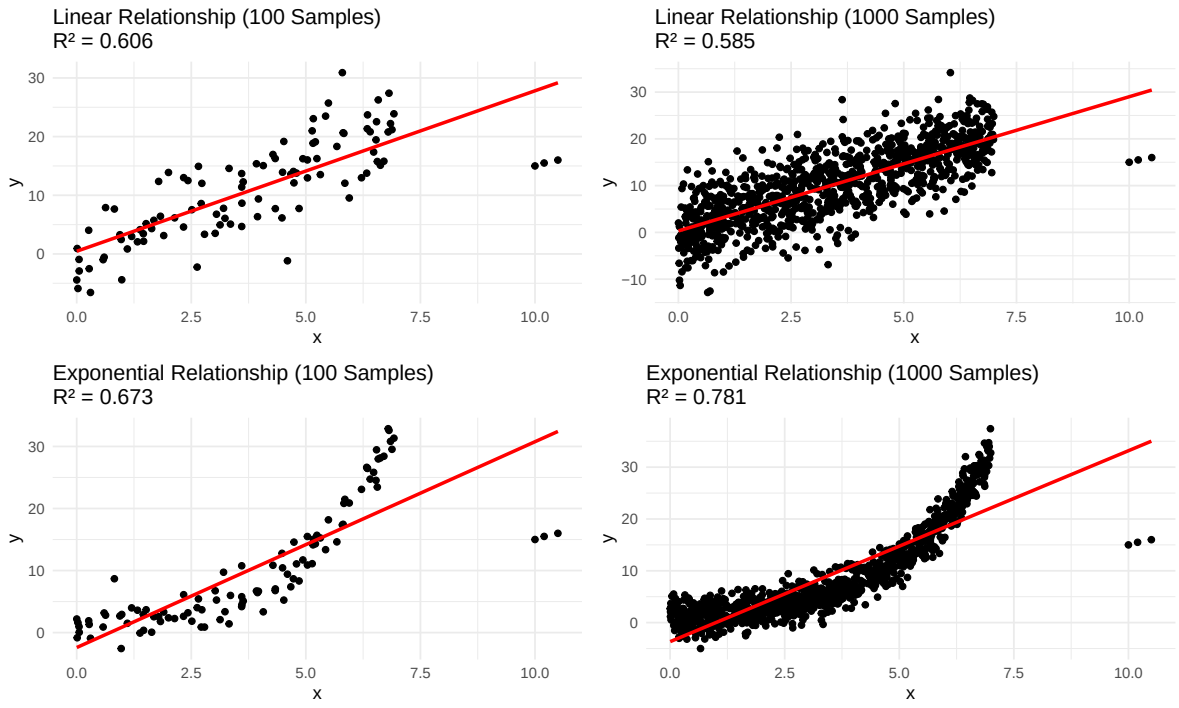

Figure S2 For HCHO, there appears to be a slight positive slope, with the difference being around 0 at SZA = 10° (JJA) and approximately 0.4 at SZA = 78° (SON). Such an SZA dependency is not observed for NO2. Is this effect negligible and statistically insignificant for HCHO, or what could be the possible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

Comments on References

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

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

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.

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

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