Supplement of Atmos. Meas. Tech., 18, 2899–2917, 2025
https://doi.org/10.5194/amt-18-2899-2025-supplement

Atmospheric Measurement Techniques
Open Access

Supplement of

Maximizing the scientific application of Pandora column observations of HCHO and NO_2

Prajjwal Rawat et al.

Correspondence to: Prajjwal Rawat (prajjwal.rawat@nasa.gov) and James H. Crawford (james.h.crawford@nasa.gov)

Supplementary materials

Supplementary Tables

Table S1. Name and location of Pandora sites used in this study.

Pandora #	Site Name	Lat, Lon (deg)	PIs
61	AldineTX	29.9011, -95.3262	Tom Hanisco
25	HoustonTX	29.7200, -95.3400	Jimmy Flynn
58	LaPorteTX	29.6721, -95.0647	Tom Hanisco
35	ManhattanNY	40.8153, -73.9505	Maria Tzortziou
38	BayonneNJ	40.6703, -74.1261	Maria Tzortziou
180	BronxNY	40.8679, -73.8781	Luke Valin
177	WestportCT	41.1183, -73.3367	Luke Valin
69	NewBrunswick	40.4622, -74.4294	Nader Abuhassan
64	NewHavenCT	41.3014, -72.9029	Nader Abuhassan
55	QueensNY	40.7361, -73.8215	Jim Szykman
134	BristolPA	40.1074, -74.8824	Luke Valin
166	Philadelphia	39.9919, -75.0811	Luke Valin
140	WashingtonDC	38.9218, -77.0124	Jim Szykman
156	HamptonVA-HU	37.0203, -76.3366	John Anderson
145	Toronto-S	43.7843, -79.1874	Vitali Fioletov

Table S2. Pandora L2 file naming for the sky-scan and direct-sun measurements and corresponding column number for vertical column, independent uncertainty, wrms, MHzD, L1, L2fit, and L2 quality flags, effective measurement time, and L1, L2fit, and L2 quality triggers.

Species	**HCHO**		**NO_2**	
Mode	**Sky-scan**	**Direct-sun**	**Sky-scan**	**Direct-sun**
Pandora L2 file naming	rfuh5p1-8	rfus5p1-8	rnvh3p1-8	rnvs3p1-8
Column # for vertical column	49	39	62	39
Column # for independent uncertainty	50	40	63	40
Column # for wrms	11	9	11	9
Column # for MHzD	51	N/A	64	N/A
Column # for L2 Quality flag	42	36	53	36
Column # for effective measurement time	3	3	3	3
Column # for L1 Quality flag	36	30	36	30
Column # for L2fit Quality flag	39	33	39	33
Column # for L1 QF DQ1 code*	37	31	37	31
Column # for L1 QF DQ2 code*	38	32	38	32
Column # for L2fit QF DQ1 code*	40	34	40	34
Column # for L2fit QF DQ2 code*	41	35	41	35
Column # for L2 QF DQ1 code*	43	37	54	37
Column # for L2 QF DQ2 code*	44	38	55	38

***Codes are given as the sum of 2^i for those data quality (DQ) parameters (i) that exceed acceptable limits.**

Supplementary Figures

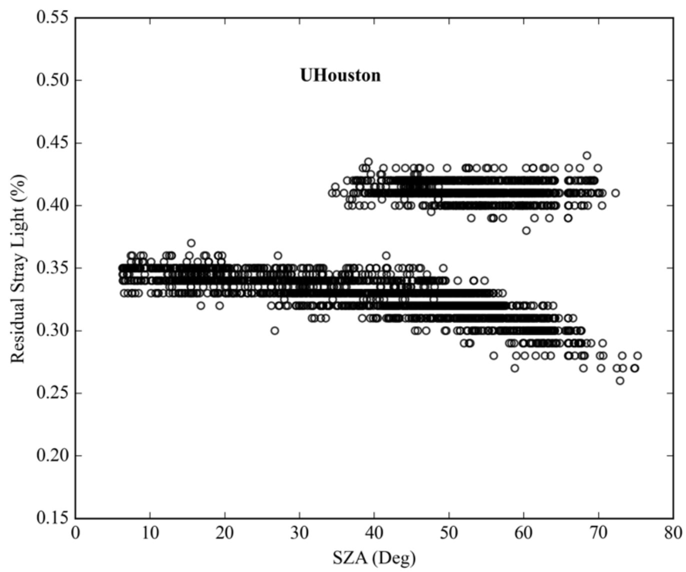

Figure S1. Residual stray light along solar zenith angle over University of Houston. Analysis was done for other sites as well and similar pattern was found for most of the sites.

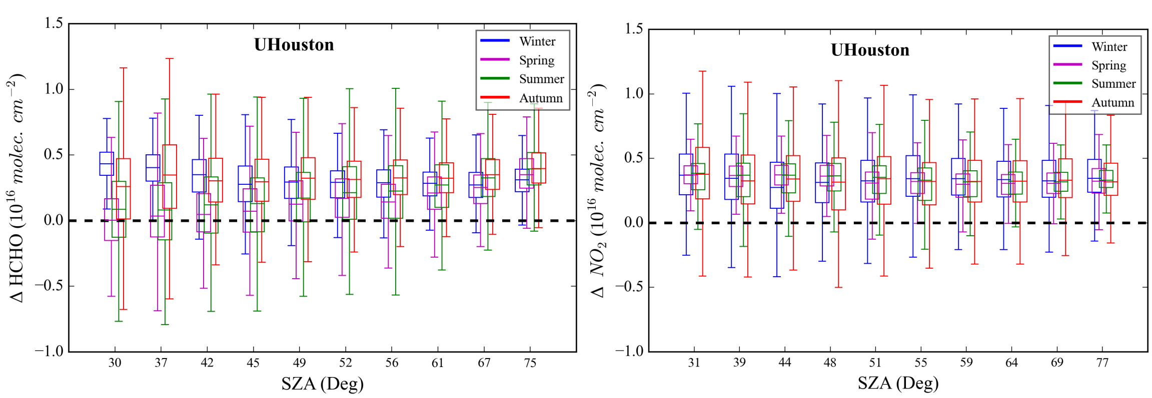

Figure S2. HCHO (left) and NO_2 (right) column differences for different seasons (DJF is winter, MAM is spring, JJA is summer, and SON is autumn) along 10 percentile bins of SZA at the University of Houston between DS and SS. Other sites are also studied, and similar results were noted.

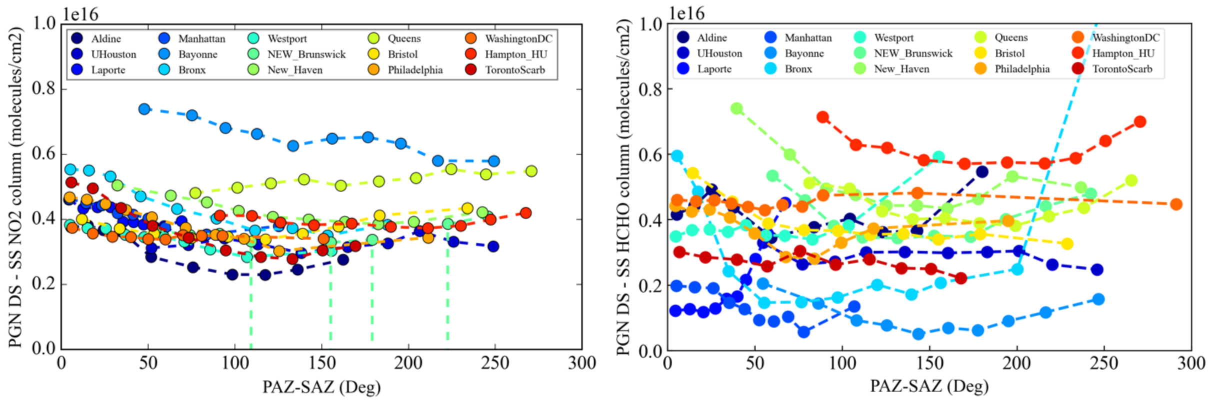

Figure S3. HCHO and NO_2 column differences along differences in pointing azimuth and solar azimuth angles (PAZ-SAZ) over 15 Pandora sites between DS and SS measurements.

Supplementary text

S1.1 Pandora instruments and data retrieval: A brief review

Software

The first operational software for Pandora instruments was called the Pan Software Suite (PSS). In 2018, the PGN team introduced the Blick Software Suite (BSS) as an evolution of PSS, incorporating various new software features (Cede et al., 2024). The main changes from PSS to BSS were real-time data, uniform file naming, quality assurance of raw and retrieved data, stray light correction, and the inclusion of sky-scan (SS) column retrievals. Currently, PGN utilizes BSSv1.8 software, released in 2020 and built upon different software iterations, including PSS and BSSv1.7. The new processor version 1.8 produces improved total column amounts using better estimations for the effective temperature and the air mass factors. Also, it adds surface concentrations and tropospheric profiles of nitrogen dioxide (NO_2) and formaldehyde (HCHO) based on sky observations and includes random and common uncertainties (Cede et al., 2024, 2025).

Hardware

Regarding the hardware, a major modification was the replacement of the front window of the telescope assembly with a window containing an anti-reflective coating (ARC) to minimize non-atmospheric spectral signal in the direct-sun (DS) observations, which is called “unwanted spectral signal” (USS) and degrades spectral fitting. This modification effectively mitigates etaloning interference and improves the ability to retrieve HCHO columns from the Pandora DS measurements (Cede et al., 2022). However, over time, the ARC tended to deteriorate, and the ultimate solution to reduce etaloning was achieved using the wedged window. The results showed that the weighted spectral fit RMS for the wedged windows were as good as for the case of no window hence wedged windows were preferred as the right optical choice for reducing unwanted spectral signal (LuftBlick_FRM4AQ, Fiducial Reference Measurements for Air Quality). Additionally, in 2019, it was observed that components made of Delrin in the head sensor were emitting formaldehyde (HCHO); subsequently, SciGlob replaced materials for these pieces from Delrin to black Nylon (Spinei et al., 2021).

Modes of operation and retrieval

In DS mode, Pandora alignment is configured such that direct beam sunlight is sampled within a narrow field of view (FOV) of 2.5° using a diffuser. For SS observations, the FOV is 1.5° without a diffuser. In the DS diffuser mode, most of the light is collected within an angle of 0.5°, which is the angle subtended by the solar disk. The spectrometer incorporates 2048×64 pixels back-thinned Hamamatsu charge-coupled device (CCD) with a 50µm entrance slit and a grating with 1200 lines mm^{-1} (Cede et al., 2022). After radiance correction, the slant column densities (SCDs) are estimated following DOAS analysis and using a static reference spectrum (taken months apart) from the modified Langley extrapolation method (MLE) (Herman et al. 2009) or extended MLE methods (in e-MLE the tropospheric SCDs are further subtracted in direct-sun using SS SCDs) (Cede et al., 2024). The subsequent retrieval of total vertical column densities (VCD) is performed by dividing the SCD by geometrical AMF for DS measurements. The primary sources of uncertainty in the DS retrieval comes from the reference spectrum and DOAS fitting error.

In SS mode, Pandora records scattered light at multiple zenith angles (0, 60, 75, 88, and 89 degrees) while maintaining a fixed azimuth direction, typically oriented northward for sites in the northern hemisphere. The Pandora SS retrieval is a fast analytical method based on rigorous background radiative transfer modeling (Cede et al., 2024; Frieß et al., 2019; Tirpitz et al., 2021), and a reference spectrum taken from zenith (0°) that is unique to each measurement. This method was referred to as the NASA algorithm during the CINDI-2 intercomparison, when the method was applied across the harmonized or median SCDs of all participating instruments. Results showed a nominal bias and strong correlation for both the HCHO and NO_2 retrievals using the NASA algorithm during clear and cloudy days, which further improved when using actual SCDs rather than the median values. Similarly, Frieß et al., (2019) also applied different MAX-DOAS retrieval algorithms, including the NASA algorithm, to the simulated SCDs with different atmospheric scenarios and reported reasonable output with minimal computational cost for the NASA analytical algorithm. In brief, the fast NASA algorithm relies primarily on a pure Rayleigh atmosphere, with the optical light path (AMF) approximated from the typical ratio of oxygen dimer (O_2-O_2) SCDs in DOAS fitting with a known O_2-O_2 vertical column from climatology, as explained by Cede et al. (2024). Generally, the derived tropospheric column is a scattered response of photons, mainly in the lower troposphere up to 4 km, and the weighted profiles are columns at various heights up to 4-5 km, but most typically in the range of 2-4 km. The exact uncertainty budget estimation for SS measurements is very challenging, and Pandora reports the measurement uncertainty for each retrieved product. However, SS has minimal reference spectrum uncertainty due to coherent zenith reference for each retrieval, unlike DS having a nearly fixed reference spectra and vice-versa for AMF.

S1.2 Pandora observational schedule

The observational schedule of Pandora is a sequence of routines to be executed by the Pandora instruments for different modes of observation. The standard schedules are stored in the BSS routine library and described by two-letter codes. Below are the codes discussed in the present manuscript. The details of all possible routines can be found on page no. 63 of Cede et al., (2024). We also give the calculation to convert t_{eff} (Column 3: Effective duration of measurement [s]) to total measurement duration (Δt) for SS routines since Pandora data files only provide a value for the maximum zenith angle.

1. Elevation scan routines E *

EK = Detailed elevation scan at standard azimuth without filter; the standard azimuth is the one given in the BlickO Instrument Configuration File, entry "Standard azimuth for elevation scans [deg]". All elevation scan routines also use entry "Maximum unobstructed pointing zenith angle at standard azimuth [deg]" from the configuration file. This routine has the following angles [0,60,70,75,"MAX-9","MAX-8","MAX-7","MAX-6","MAX-5","MAX-4","MAX-3","MAX-2","MAX-1","MAX","MAX-1","MAX-2","MAX-3","MAX-4","MAX-5","MAX-6","MAX-7","MAX-8","MAX-9",75,70,60,0], where for a flat horizon MAX angle is set to 89.0 deg. It could also be larger or smaller than 89.0 deg, depending on the horizon of the location at the standard azimuth. Total measurement duration can be calculated as $\Delta t = 26^* t_{eff} /2 + t_{eff}$.

EL = Detailed elevation scan at standard azimuth with U340 filter. This routine has the following angles [0,40,50,60,70,75,80,82,85,"MAX-2","MAX-1","MAX","MAX-1","MAX-2",85,82,80,75,70,60,50,40,0], where for a flat horizon MAX angle is set to 89.0 deg. It could also be larger or smaller than 89.0 deg, depending on the horizon of the location at the standard azimuth. Total measurement duration can be calculated as $\Delta t = 22^* t_{eff} /2 + t_{eff}$.

EO = Quick elevation scan (5 zenith angles) at standard azimuth without filter. This routine has the following angles [0,60,75,"MAX-1","MAX","MAX-1",75,60,0], where for a flat horizon MAX angle is set to 89.0 deg. It could also be larger or smaller than 89.0 deg, depending on the horizon of the location at the standard azimuth. Total measurement duration can be calculated as $\Delta t = 8^* t_{eff} /2 + t_{eff}$.

EU = Quick elevation (5 zenith angles) at standard azimuth with U340 filter. This routine has the following angles [0,60,75,"MAX-1","MAX","MAX-1",75,60,0], where for a flat horizon MAX angle is set to 89.0 deg. It could also be larger or smaller than 89.0 deg, depending on the horizon of the location at the standard azimuth. Total measurement duration can be calculated as $\Delta t = 8^* t_{eff} /2 + t_{eff}$.

Find routines F*

FS = Find Sun; long sun search for all spectrometers, saves final figure and averaged data, but not spectral data

FD = FS, but with diffuser in the optical path

FO = FD but scanning only around the center of the FOV with longer measurement duration. This is not used to "find" the sun.

FP = FU but scanning only around the center of the FOV with longer measurement duration. This is not used to "find" the sun.

FU = FS, but with U340 filter plus diffuser combination in the optical path

FW = FS with spectra, but showing the results in 4 wavelength regions

2. **Direct-sun routines S***

SQ = Quick direct-Sun observation without filter

SS = Quick direct-Sun observation with U340 filter

3. **Zenith sky routines Z***

ZO = Zenith-Sky observation without filter

ZU = Zenith-Sky observation with U340 filter

S1.3 Main causes of low- and medium-quality flags and characterization of recovered data

S1.3.1 Data flagging procedure

PGN quality flags are assigned in various stages from L1 (raw data) to L2fit (spectral fit data) to L2 (processed data) for ensuring data quality through multiple checks at each stage. At the L1 stage, data is flagged into either low- or medium-quality based on instrument-related issues such as excessive dark counts, detector saturation, dark count differs significantly from the dark map for too many pixels, different effective temperature, and unsuccessful dark background fitting. In the L2fit stage, where spectral fitting is performed, data is further flagged based on factors including the quality of the fit, the wrms limit (normalized rms of fitting residuals weighted with independent uncertainty), and wavelength shift. Finally, at the final retrieval L2 stage, factors including retrieval error and atmospheric variability are used to flag data into the medium or low-quality. The medium and low-quality threshold values for each data quality indicator (DQ) are estimated using the Gaussian mixture regression model and are described in the Fiducial Reference Measurements for Air Quality report of LuftBlick for each retrieved parameter (Gebetsberger et al., 2022). This step provides quality flags as high (10), medium (11), and low (12) quality data without quality assurance. In the second quality assurance step, the time series fluctuations of the quality indicators are studied for a more extended period and sudden relative changes or drifts are used to categorize the data in high (0), medium (1), and low (2) quality 'with quality assurance'. However, these steps do not account for various other potential sources of errors in the data, such as misalignment of instruments, telescope collimator errors, or unknown spectral features. To address such issues, a third quality assurance procedure is implemented based on time series fluctuation of multiple parameters of the multiple data products and assigns invalid/unusable data flags as high unusable (20), medium unusable (21), and low unusable (22), data if anomalous behavior is observed. This is discussed in more detail by Gebetsberger et al. (2022). The second and third steps may be performed over a longer time period than the first step, and it is thus possible for newer sites to have no data categorized as 0, 1, 2 or 20, 21, 22.

S1.3.2 Propagation of the Pandora quality flags from L1, to L2Fit, to L2 for high, medium, and low-quality data for the University of Houston site.

Here we provide the data quality processing workflow for the data from Figures 6 to 9 in the main text for the University of Houston Pandora instrument. The present analysis traces the assignment of quality flags (QFs) from L1 to L2, using only L2 QFs of 0, 1, 2, 10, 11, 12. No analysis has been performed for unusable data (QFs 20, 21, and 22).

Direct-sun HCHO

Figure S4a shows the contribution of different L1 data QFs for HCHO direct-sun at the University of Houston Pandora. A total of 82% of data is high quality unassured (QF=10), while 15% is low quality unassured (QF=12), and 3% is medium quality unassured (QF=11). Table S3 shows the reasons for flagging the points in Figure S4 as medium or low-quality. Most of the medium quality data is flagged in the L1 step due to high dark counts (936 points trigger L1 code #8) and "dark count differs significantly from the dark map" (136 points trigger L1 code #32). The main trigger for low quality data in the L1 step is saturated data (5855 data points trigger L1 code #1).

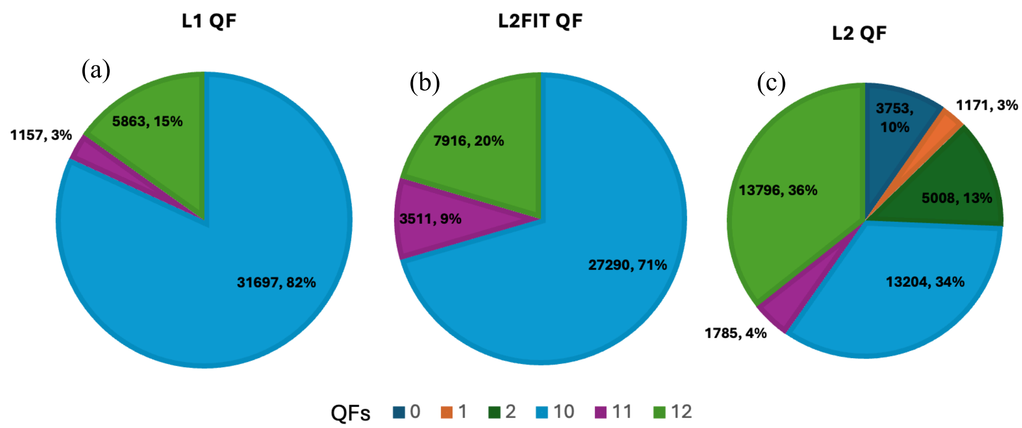

Figure S4. Contribution of Pandora (a) L1, (b) L2fit, and (c) L2 data QFs for HCHO direct-sun at the University of Houston Pandora (PGN #25). The total and percent number of points in each flag is provided for each step.

Figure S4b shows the contribution of different L2fit data QFs. 11% of the data is shifted from high quality (QF=10) to low or medium quality. This shift is due "Normalized rms of fitting residuals weighted with independent uncertainty too large" (3233 points trigger L2fit code #8 and 9 for medium quality, 5022 points trigger L2fit code #8 and 9 for low quality). This corresponds to the wrms discussed in the main text and is hereafter referred to as such. The wrms cutoff is 4.9×10^{-4} for medium-quality data and 7.0×10^{-4} for low-quality data in PGN (Cede et al., 2025).

Figure S4c shows the final L2 QFs. Here there is a shift of 27% of the data from high quality unassured (QF=10) to lower QFs. This results in only 44% high quality data (QF=10 or 0) compared to 82% in the first step (Figure S4a). 49% of the data is now low quality (QF=12 or 2) and 8% is medium quality (QF=11 or 1). Additional medium and low QF points in the L2 stage are triggered mainly due to high atmospheric variability (1044 points trigger L2 code #8 and 9 for medium quality, 17446 points for low quality). These triggers correspond to the standard PGN cutoffs for atmospheric variability of 61% for medium quality and 70% for low quality (Cede et al., 2025).

The main reason for the loss of data at this site is therefore 50% due to large atmospheric variability, 21% due to large wrms, and 28% due to saturated data, with a small contribution from dark counts.

Data flags in the data rescued using the methodology described in the main text (Figure 10).

After following the methodology described in Figure 10 in the main text for ensuring Pandora data quality and availability for scientific applications for the University of Houston site, it was possible to determine the reasons that restored data was flagged as medium or low quality. In Table S3, the red data in the 'Usable' column shows the amount of restored data in each category. In the L1 QF step, 33% of the data (2354 points) flagged as QF=11 or 12 were recovered, primarily due to recovering points with high dark counts (L1 code #8) and saturated data (L1 code #1). In the L2fit stage, 40% of the flagged data (4547 points) was recovered primarily under the code for high wrms (L2fit code #8). In the L2 stage, 72% (15595 points) of the data were recovered using the new method, with 13063 points primarily restored from those flagged for high atmospheric variability (L2 code #8). This finding suggests that the atmospheric variability, wrms, and dark count and saturated data threshold for direct-sun HCHO are too stringent in the standard PGN quality flags.

Table S3. Data flagging codes for Pandora direct-sun HCHO at the University of Houston Pandora site. Pandora data exceeding the DQ1 (medium quality) and DQ2 (low quality) limits for each indicator code (sum over 2^i)are provided., The recovered data deemed usable are shown in red. The total number of unflagged and flagged points in each step (L1, L2fit, L2) is 38717 (see Figure S4).

Code No	Indicator	L1 QF				Indicator	L2fit QF				Indicator	L2 QF			
		DQ1		DQ2			DQ1		DQ2			DQ1		DQ2	
		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable
1	Saturated data			5855	1615	L1 data quality above 0	278	254	2886	1441	L2Fit data quality above 0	1834	1833	1276	562
2	Too few dark counts			1	1	Spectral fitting was not successful					Retrieval error				
4	No temperature given or effective T diff	43	43			Wavelength shift too large					Air mass factor too large				
5											4+1	72	72	69	54
8	Dark count high	936	645			Normalized rms of fitting residuals weighted with independent uncertainty too large	2929	1546	2049	874	Atmospheric variability too large	940	939	10888	10151
9						8+1	304	257	2973	175	8+1	104	104	6558	1869
10						8+2			1	0					
12						8+4			3						
13						8+4+1			4		8+4+1	6	6	13	5

Code No	Indicator	L1 QF				Indicator	L2fit QF				Indicator	L2 QF			
		DQ1		DQ2			DQ1		DQ2			DQ1		DQ2	
		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable
16	Unsuccessful dark background fitting	1	1												
32	The dark count differs significantly from the dark map for too many pixels	136	34	3	0										
33	32+1			4	0										
40	32+8	41	15												

Sky-scan HCHO

Figure S5a shows the contribution of different L1 data quality flags for HCHO sky-scan at the University of Houston Pandora. 53% of data is high quality (QF=10 or 0), 34% is medium quality (QF=11 or 1), and 12% is low-quality (QF=12 or 2). Table S4 shows the reasons for flagging the points in Figure S5 into medium- and low-quality flags. In the L1 flagging stage, 53% of the medium quality data is being triggered due to high dark counts (10533 points trigger L1 code # 8), and 39% by the dark count differs significantly from the dark map for too many pixels (7843 points trigger L1 code #32). Data is mainly triggered to low quality due to saturated data (5173 points trigger L1 code #1).

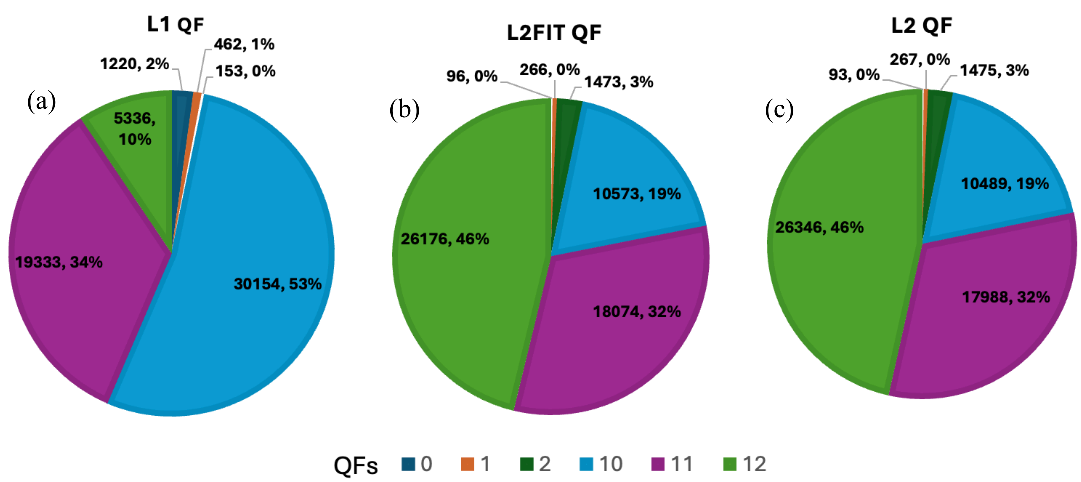

Figure S5. Contribution of Pandora (a) L1, (b) L2fit, and (c) L2 data QFs for HCHO sky-scan at the University of Houston Pandora (PGN #25). The total and percent number of points in each flag is provided for each step.

Figure S5b shows data shifts from 53% high quality to 19% high quality in L2fit stage with 32% medium quality and 46% low quality. This shows that most of the data is flagged as low and medium quality in the L2fit flagging stage. Additional medium and low QF points in the L2fit stage are triggered mainly due to high wrms (15687 points trigger L2 code #8 and 9 for medium quality, 24117 points for low quality). This corresponds to the standard PGN wrms threshold of 1.12×10^{-3} for medium quality and 1.5×10^{-3} for low quality (Cede at el., 2025). Also as shown in Section 2.2 in the main text, HCHO is observed with an U340 and open filter. In the L1 stage, if the filter wheel position is 2 (with U340 filter), about 62% of data is flagged as medium quality, while if the filter wheel position is 1 (with open filter), about 76% of data is high quality in L1 stage. That shows in the SS HCHO retrieval using open and U340 filter has some dependence of filter wheel in the data quality indicator (dark counts).

Figure S5c shows the contribution of different L2 data quality flags. Only a small amount (<1%) of additional data is flagged as low or medium quality due to high atmospheric variability (L2 code #8). The main reason for the loss of HCHO sky-scan data at this site is therefore 58% due to high wrms in the L2fit stage, and 40% due to high dark counts, dark count differs significantly from the dark map for too many pixel, and saturated data in the L1 stage.

Data flags in the data rescued using the methodology described in the main text (Figure 10).

In Table S4 the red column shows the usable data from each quality indicator. In the L1 step, 75% of the data (18122 points) flagged as QF=11 or 12 were recovered, primarily for points with high dark counts (L1 code #8), the dark count differs significantly from the dark map for too many pixels (L1 code #32), and saturated data (L1 code #1). In the L2fit stage, approximately 70% of the flagged data (24054 points) were recovered, which were flagged due to high wrms (L2fit code #8).

Table S4. Same as Table S3 for sky-scan HCHO. The total number of unflagged and flagged points in each step (L1, L2fit, L2) is 56658 (see Figure S5).

Code No	Indicator	L1 QF				Indicator	L2fit QF				Indicator	L2 QF			
		DQ1		DQ2			DQ1		DQ2			DQ1		DQ2	
		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable
1	Saturated data			5173	4014	L1 data quality above 0	2645	2623	3277	3151	L2Fit data quality above 0	18146	17776	27035	16224
2	Too few dark counts					Spectral fitting was not successful	4	4			Retrieval error				
4	No temperature given or effective T diff	38	37	1	1	Wavelength shift too large					Air mass factor too large				
8	Dark count high	10533	8424	3	2	Normalized rms of fitting residuals weighted with independent uncertainty too large	11510	11217	22565	12842	Atmospheric variability too large	38	36	172	161
9				1	0	8+1	4177	4104	1552	511	8+1	71	63	614	316
10						8+2			3	0					
11						8+2+1			2	0					
12		2	1			8+4	3	3	247	36					

Code No	Indicator	L1 QF				Indicator	L2fit QF				Indicator	L2 QF			
		DQ1		DQ2			DQ1		DQ2			DQ1		DQ2	
		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable
13						8+4+1	1	1	3						
16	Unsuccessful dark background fitting	201	141												
24	16+8	27	21												
28	16+8+4	1	1												
32	The dark count differs significantly from the dark map for too many pixels	7843	5295	283	185										
36	32+4	6	6												
40	32+8	1124	853												
256	Absolute value of retrieved wavelength shift too large	13	0	2	0										
768	512+256	7	0	26	0										

NO_2 direct-sun

Figure S6a shows the contribution of different L1 data quality flags for NO_2 direct-sun over the University of Houston Pandora. Most of the data passes L1 data quality filtering (96%). Table S5 shows the reasons for flagging the points in Figure S6 into medium and low-quality flags.

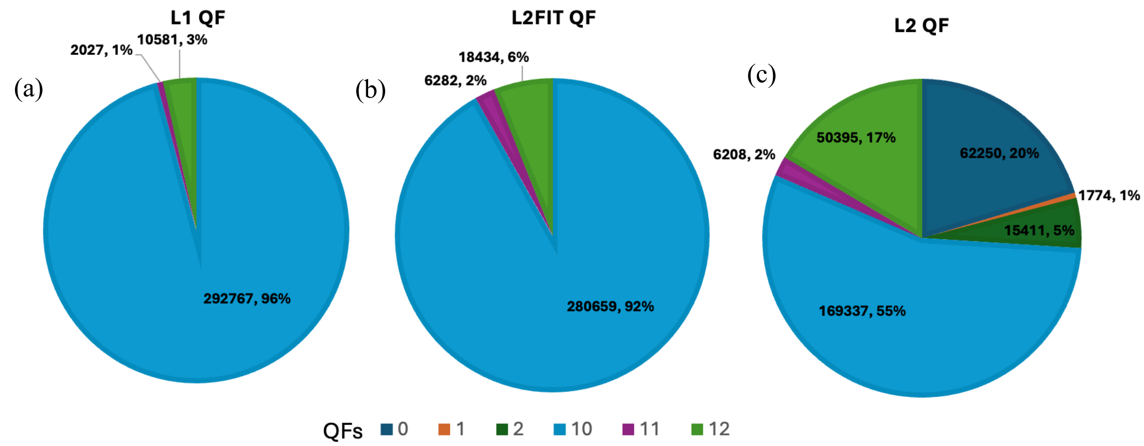

Figure S6. Contribution of Pandora (a) L1, (b) L2fit, and (c) L2 data quality for NO2 direct Sun over UHouston Pandora (PGN #25). The quality flags are written outside over the pie chart and their percentage with total number of points inside the sector of pie chart.

Figure S6b shows the contribution of different L2fit data quality flags. Only about 4% of high-quality data in the L1 stage is now flagged as medium or low quality in the L2fit stage, in both cases due to high wrms (5559 points triggered to medium quality and 10493 points triggered to low quality by L2fit code #8 and 9). Figure S6c shows the contribution of different L2 data quality flags. High quality data is reduced from 92% to 75%. The main trigger of medium- and low-quality data is high atmospheric variability (4147 points trigger to medium quality and 59225 points triggered to low quality by L2 code #8 and 9).

Overall, for NO_2 direct-sun, approximately 16% of data is flagged due to exceeding the PGN atmospheric variability thresholds of 36% and 42% (Cede et al., 2025), 4% is flagged due to high wrms, with the remainder exceeding flags in the L1 stage including issues with temperature and dark counts.

Data flags in the data rescued using the methodology described in the main text (Figure 10).

After following the procedure in Figure 10, we find that in the L1 stage, over 90% of the data (2354 points) flagged as QF=11 or 12 is usable, which are flagged primarily due to high dark counts (L1 code #8) and saturated data (L1 code #1). In the L2fit stage, 80% of the low- or medium-quality data flagged with high wrms is recovered (L2fit code #8). In the L2 stage, almost all the data flagged for high atmospheric variability are recovered (51514 points under L2 code #8).

Table S5. Same as Table S3 but for direct-sun NO_2. The total number of unflagged and flagged points in each step (L1, L2fit, L2) is 305375 (see Figure S6).

Code No	Indicator	L1 QF				Indicator	L2fit QF				Indicator	L2 QF			
		DQ1		DQ2			DQ1		DQ2			DQ1		DQ2	
		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable
1	Saturated data			9977	9420	L1 data quality above 0	723	723	7899	7877	L2Fit data quality above 0	1830	1824	6442	4523
2	Too few dark counts			16	16	Spectral fitting was not successful					Retrieval error				
4	No temperature given or effective T diff	480	480	10	10	Wavelength shift too large					Air mass factor too large	783	783		
5	4+1											1201	1201	26	11
8	Dark count high	70	62			Normalized rms of fitting residuals weighted with independent uncertainty too large	5512	5458	7819	5415	Atmospheric variability too large	4030	4030	47528	47484
9						8+1	47	47	2674	2139		117	113	11697	10708
12						8+4			34			10	10		
13						8+4+1			8	6		11	11	113	0
16	Unsuccessful dark	14	14												

Code No	Indicator	L1 QF				Indicator	L2fit QF				Indicator	L2 QF			
		DQ1		DQ2			DQ1		DQ2			DQ1		DQ2	
		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable
	background fitting														
32	The dark count differs significantly from the dark map for too many pixels	1399	1385	569	569										
33	32+1			9	7										
40	32+8	64	62												

NO_2 sky-scan

Figure S7a shows the contribution of different L1 data quality flags for NO_2 sky-scan at the University of Houston Pandora. Table S6 shows the medium and low-quality trigger for data points in Figure S7. A total of 79% of data is high quality (QF=10 or 0), while 9% is medium quality (QF=11 or 1), and 12% is low quality (QF=12 or 2). In the L1 flagging stage, medium quality data is mainly due to dark count differs significantly from the dark map for too many pixels (2385 points trigger L1 code #32). Low quality data is mainly triggered by saturated data (3612 points trigger L1 code #1).

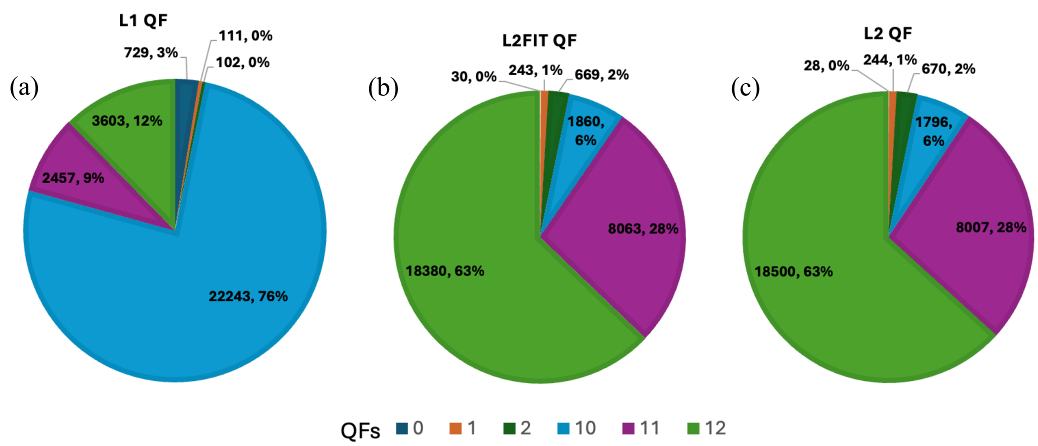

Figure S7. Contribution of Pandora (a) L1, (b) L2fit, and (c) L2 data quality for NO2 sky scan over UHouston Pandora (PGN #25). The quality flags are written outside over the pei chart and their percentage with total number of points inside the sector of pie chart.

Figure S7b shows the contribution of different L2fit data quality flags. High quality data is reduced from 79% to only 6%. The shift to both medium and low-quality data is due to high wrms (8305 points trigger L2 code #8 and 9 for medium quality, 16841 points for low quality trigger L2 code #8 and 9). This corresponds to the standard PGN threshold for wrms of 6.9×10^{-4} for medium quality and 9.1×10^{-4} for low quality (Cede et al., 2025). Less than 1% of the data is further flagged in the L2 stage (Figure S7c) due to high atmospheric variability (L2 code #8).

The main reason for the loss of data at this site is therefore 80% due to high wrms in the L2fit stage, and 20% due to dark count differs significantly from the dark map for too many pixels and saturated data in the L1 stage.

Data flags in the data rescued using the methodology described in the main text (Figure 10).

In Table S6, the red column shows the usable data from each quality indicator issues starting from L1 to L2fit to L2 stage for sky-scan NO_2. In the L1 step, 60% of the data (18122 points) flagged as QF=11 or 12 were recovered, primarily due to recovering points with issues with saturated data and dark counts difference (L1 codes #1 and #32). In the L2fit stage, more than 95% of the flagged data was recovered primarily under the code for high wrms (L2fit code #8). In the L2 stage, nearly all data flagged for high atmospheric variability (L2 code #8) was recovered although this was <1% of the data.

Table S6. Same as Table S but for sky-scan NO_2. The total number of unflagged and flagged points in each step (L1, L2fit, L2) is 29245 (see Figure S7).

Code No	Indicator	L1 QF				Indicator	L2fit QF				Indicator	L2 QF			
		DQ1		DQ2			DQ1		DQ2			DQ1		DQ2	
		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable
1	Saturated data			3612	3544	L1 data quality above 0			1998	1997	L2Fit data quality above 0	8175	8165	18618	18226
2	Too few dark counts					Spectral fitting was not successful					Retrieval error				
4	No temperature given or effective T diff	26	26	1	1	Wavelength shift too large					Air mass factor too large				
8	Dark count high	63	63			Normalized rms of fitting residuals weighted with independent uncertainty too large	8262	8251	15703	15493	Atmospheric variability too large	26	26	121	120
9						8+1	43	43	1138	1076	8+1	50	50	431	410
10						8+2	1	1							
12						8+4			195	68					
13						8+4+1			15	2					

Code No	Indicator	L1 QF				Indicator	L2fit QF				Indicator	L2 QF			
		DQ1		DQ2			DQ1		DQ2			DQ1		DQ2	
		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable		PGN	Usable	PGN	Usable
16	Unsuccessful dark background fitting	75	74												
32	The dark count differs significantly from the dark map for too many pixels	2385	2331	79	72										
40	32+8	12	11												
256	Absolute value of retrieved wavelength shift too large	4	0	2	0										
768	512+256	3	0	11	0										