

FRM4GHG 2.0

Fiducial Reference Measurements for Greenhouse Gases



D2.5.1 TN: AirCore retrieval procedure updates

Deliverable:	D2.5.1
Date:	13/10/2026
Lead authors:	Huilin Chen, Xinyu Yang, Steven van Heuven, and Rigel Kivi
Subject:	ESA-Contract No. 4000136108/21/I-DT-lr
Category:	ESA Express Procurement (EXPRO)
Our ref.:	Proposal, FRM4GHG 2.0 – Expro, ESA RFP/3-17031/21/I-DT-lr, revision from 27 August 2021 and 7 September 2021

Table of contents

1	Document change record.....	3
2	Access list.....	3
3	Purpose	3
4	Document structure	3
5	Results.....	3
6	OUTLOOK.....	26
7	Applicable documents.....	26
8	Reference documents.....	26

1 DOCUMENT CHANGE RECORD

Issue	Date	Item	Comment
V1.0	2026-08-29	D2.5.1	Initial version

2 ACCESS LIST

This document is a deliverable “D2.5.1: AirCore retrieval procedure updates created for the project FRM4GHG 2.0 and will be submitted to ESA. The document can be downloaded from the project webpage <http://frm4ghg.aeronomie.be>.

3 PURPOSE

The document describes the updated AirCore retrieval procedure, which aims to improve the accuracy of the retrieved AirCore profiles at Sodankylä, especially in terms of altitude registration. The improved AirCore retrieval procedure will also allow to create a new version of Sodankylä AirCore profiles.

4 DOCUMENT STRUCTURE

See Table of Contents

5 RESULTS

WP 2: Improvements to retrieval algorithms for GHG and additional species from low-resolution and TCCON spectrometers

5.1 Description of AirCore Dataset V2.0

The updated AirCore Dataset V2.0 consists of reprocessed AirCore vertical profiles measured at Sodankylä using the improved retrieval procedure described in this study. The new retrieval framework primarily aims to improve the altitude registration accuracy of the retrieved AirCore profiles, particularly in the stratosphere, where pressure disequilibrium during sampling can introduce significant altitude attribution biases.

Compared with the conventional AirCore retrieval procedure, AirCore retrieval procedure V2.0 mainly introduces a revised altitude registration scheme based on an updated starting-point retrieval

method. The new approach was developed and evaluated using six flights with simultaneous AirCore and LISA observations, including Flight 20170426, Flight 20170904, Flight 20170905, Flight 20170906, Flight 20180619, and Flight 20180625. For these flights, a series of AirCore profiles were generated by progressively shifting the retrieval starting point. Each retrieved profile was then compared with the corresponding LISA observations to determine the starting-point range that minimized the weighted RMSE. The optimized starting points derived from the six flights were subsequently used to establish the representative retrieval framework for Dataset V2.0.

In the original AirCore V1.0 retrieval procedure, the altitude attribution assumed instantaneous pressure equilibrium between the inside and outside of the AirCore tube during sampling. Under this assumption, the retrieval starting point was defined as the midpoint of the fill gas transition region, corresponding to the 50% position between residual fill gas and atmospheric sample air (Fig. 1, left panel, Scheme A). This transition region represents the center of the mixing zone between residual fill gas and atmospheric air.

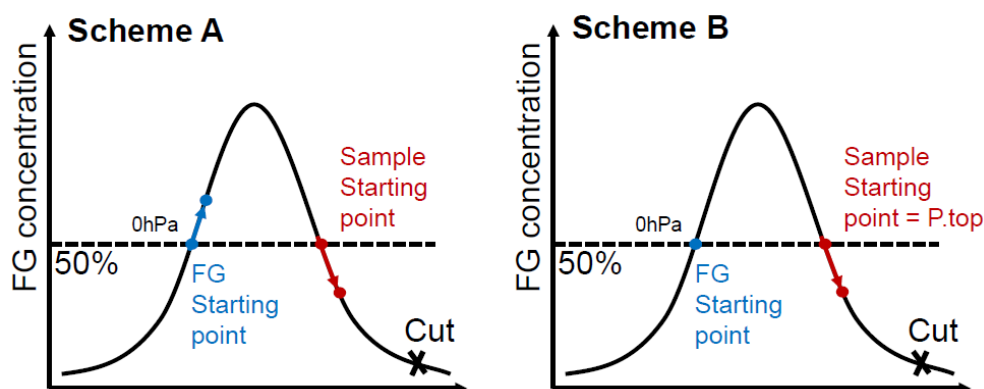


Figure 1. Original retrieval scheme vs. starting point optimization scheme

However, residual fill gas inside the AirCore tube does not participate in the atmospheric gas flow process during the balloon descent and therefore should not be shifted together with the atmospheric sample region. In retrieval procedure V2.0, the residual fill gas reference point is kept fixed at the original 50% fill gas position, while only the atmospheric sample starting point is shifted during the retrieval procedure. In addition, the pressure corresponding to the sample starting point is initialized using the balloon top pressure measured by the AirCore data logger (Fig. 1, right panel, Scheme B). Using this revised retrieval framework, the six validation flights yielded an average optimized starting point of $20\% \pm 4\%$ (Fig. 2, Starting point vs. RMSE). The optimized starting point for each individual flight is summarized in Table 1.

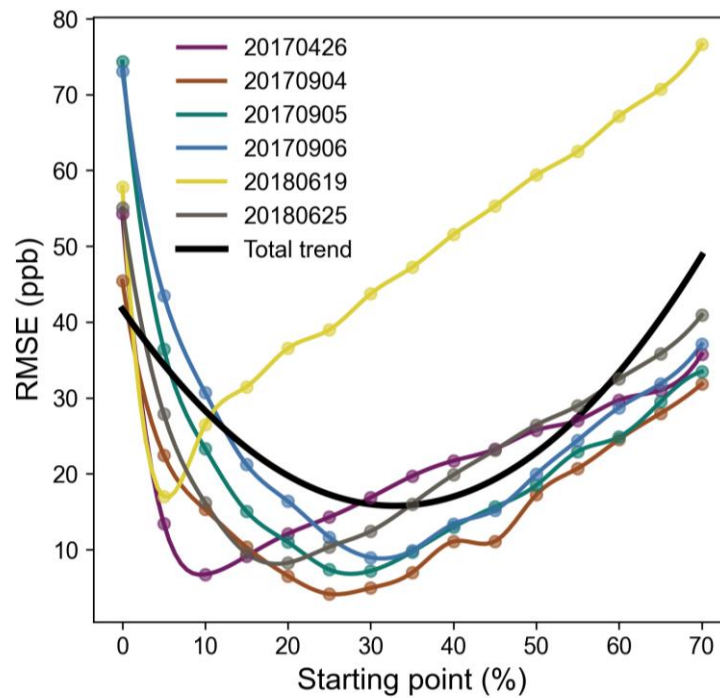


Figure 2. The RMSE varies with the starting point in the optimization scheme

Table 1. The optimal starting point of each flight

Flight	20170426	20170904	20170905	20170906	20170619	20170625
Optimized starting point	10%	25%	30%	30%	5%	20%

During AirCore analysis, the gas sequence inside the measurement system is typically: fill gas → calibration gas → residual fill gas → atmospheric sample air. The amount and duration of residual fill gas directly influence the effective location of the sample starting point. To further constrain the starting-point determination, an additional diagnostic method was developed based on the amount of the residual fill gas. The ratio between the fill gas time interval and the total AirCore analysis time was used together with the surface pressure to estimate a diagnosed top pressure. The relationship between the difference of diagnosed top pressure and the actual balloon top pressure measured by the AirCore data logger was then used to determine the retrieval starting point for each flight (Fig. 3, Scheme E conceptual diagram).

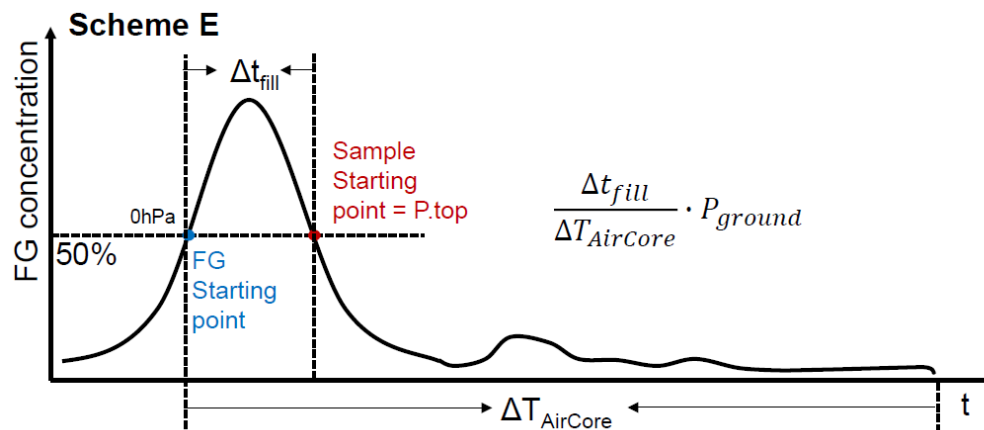


Figure 3. Scheme for fill gas time derived starting point

The empirical fitting relationship between pressure difference (diagnosed top pressure - actual balloon top pressure) and the optimized retrieval starting point by AirCore-LISA comparison is shown in Fig. 4 (left panel), with an R^2 value of 0.8 and a linear fit of: $y = 16.23x + 62.6$. Only five flights were used in this fitting analysis because Flight 20170426 was excluded due to malfunction of the AirCore data logger, which prevented an accurate determination of the balloon top pressure. To avoid extrapolating flights with small pressure differences to unrealistic starting points, an equilibrium point (0 hPa, 50% starting point) was further introduced as a constraint for the linear fitting. This modification produced an improved fitting relationship with a higher correlation coefficient ($R^2 = 0.9$): $y = 12.2x + 52.1$ (Fig. 4, right panel). This relationship works in the range when the difference between diagnosed top pressure and actual balloon top pressure is less than 0 hPa.

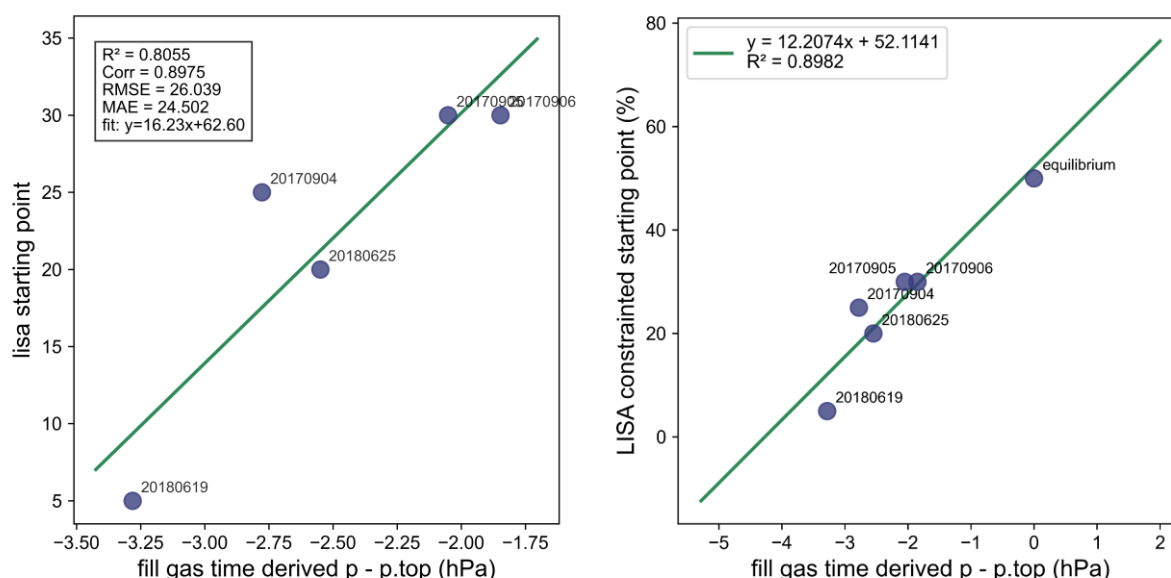


Figure 4. Linear relationship between fill gas time derived top pressure and LISA constrained starting point

Using the final diagnostic fitting relationship, retrieval starting points were recalculated for all AirCore flights measured during 2013–2025. In case the pressure difference between diagnosed top pressure and actual balloon top pressure is larger than 0 hPa, a default 25% was used. The diagnosed starting points and corresponding retrieval information for all flights are summarized in Table 2. The resulting AirCore Dataset V2.0 provides updated pressure-altitude registration and improved vertical trace gas profiles for CO_2 , CH_4 , and CO measurements, thereby establishing a more physically consistent dataset for studies of stratospheric composition, transport processes, and long-term changes in the mean age of air.

Table 2. Fill gas time derived optimal starting point of AirCore flights (2013-2025)

Year	Date	Number	composition	Data describe	Starting point
2013	20130903	1-2	CO2, CH4, CO	No data	/
	20130905	3-4	CO2, CH4, CO	read failed	/
	20131022	5	CO2, CH4, CO	Processed	25
2014	20140319	8	CO2, CH4, CO	Processed	25
	20140409	9	CO2, CH4, CO	Processed	25
	20140508	10	CO2, CH4, CO	Processed	25
	20140714	11	CO2, CH4, CO	Processed	25
	20140715	12	CO2, CH4, CO	Processed	25
	20140716	13	CO2, CH4, CO	Processed	25
	20140717	14	CO2, CH4, CO	datalogger loss	25
	20140718	15	CO2, CH4, CO	datalogger loss	25
2015	20150620	16	CO2, CH4, CO	Processed	25
	20150621	17	CO2, CH4, CO	datalogger loss	/
	20150622	18	CO2, CH4, CO	read failed	/
	20150623	19	CO2, CH4, CO	Processed	26
2016	20160204	20	CO2, CH4, CO	Processed	25
	20160310	21	CO2, CH4, CO	Processed	22
	20160414	22	CO2, CH4, CO	Processed	48
	20160526	23	CO2, CH4, CO	No data	/
	20160712	24	CO2, CH4, CO	Processed	38
	20160824	25	CO2, CH4, CO	Processed	48
	20160829	26	CO2, CH4, CO	Radiosonde's descent phase data is partially missing.Parachut did not work well, crashed	/
2017	20170421	27	CO2, CH4, CO	Processed	50
	20170424	28	CO2, CH4, CO	Processed	25
	20170426	29	CO2, CH4, CO	datalogger loss, valve close manually	/
	20170515	30	CO2, CH4, CO	Processed	46
	20170828	31	CO2, CH4, CO	Processed	25
	20170904	32	CO2, CH4, CO	Processed	26
	20170905	33	CO2, CH4, CO	Processed	32
	20170906	34	CO2, CH4, CO	Processed	25
	20170907	35	CO2, CH4, CO	Processed	31
2018	20171009	36	CO2, CH4, CO	Processed	30
	20180417	37	CO2, CH4, CO	Processed	48
	20180528	38	CO2, CH4, CO	Processed	34
	20180618	39	CO2, CH4, CO	Processed	25
	20180619	40	CO2, CH4, CO	Processed	14
	20180620	41	CO2, CH4, CO	Processed	40
	20180625	42	CO2, CH4, CO	Processed	28
	20180702	43	CO2, CH4, CO	Processed	38
	20180801	44	CO2, CH4, CO	Radiosonde pressure values missing before landing	45
	20181003	45	CO2, CH4, CO	Processed	25

2019	20190410	46 CO ₂ , CH ₄ , CO	Processed	39
	20190628	47 CO ₂ , CH ₄ , CO	Processed	35
	20190724	48 CO ₂ , CH ₄ , CO	radiosonde loss	/
	20190801	49 CO ₂ , CH ₄ , CO	Processed	42
	20190828	50 CO ₂ , CH ₄ , CO	Processed	48
	20190909	51 CO ₂ , CH ₄ , CO	Processed	29
2020	20200617	52 CO ₂ , CH ₄ , CO	Processed	25
	20200819	53 CO ₂ , CH ₄ , CO	Processed	28
2021	20210419	54 CO ₂ , CH ₄ , CO	Processed	30
	20210426	55 CO ₂ , CH ₄ , CO	Processed	37
	20210817	56 CO ₂ , CH ₄ , CO	Processed	26
	20210822	57 CO ₂ , CH ₄ , CO	Processed	32
	20210823	58 CO ₂ , CH ₄ , CO	Processed	44
	20210826	59 CO ₂ , CH ₄ , CO	Processed	32
2022	20220331	60 CO ₂ , CH ₄ , CO	Processed	31
	20220505	61 CO ₂ , CH ₄ , CO	Processed	40
	20220629	62 CO ₂ , CH ₄ , CO	Processed	34
	20220825	63 CO ₂ , CH ₄ , CO	Processed	42
2023	20230427	64 CO ₂ , CH ₄ , CO	Processed	50
	20230808	65 CO ₂ , CH ₄ , CO	Processed	/
	20230818	66 CO ₂ , CH ₄ , CO	Processed	50
	20230915	67 CO ₂ , CH ₄ , CO	Processed	50
2024	20240418	68 CO ₂ , CH ₄ , CO	Processed	47
	20240902	69 CO ₂ , CH ₄ , CO	Processed	25
2025	20250312	70 CO ₂ , CH ₄ , CO	Processed	25
	20250701	71 CO ₂ , CH ₄ , CO	Processed	46
	20250703	72 CO ₂ , CH ₄ , CO	Processed	25
	20250829	73 CO ₂ , CH ₄ , CO	Processed	25
	20250830	74 CO ₂ , CH ₄ , CO	Processed	25
	20250831	75 CO ₂ , CH ₄ , CO	Processed	25
	20250902	76 CO ₂ , CH ₄ , CO	Processed	25

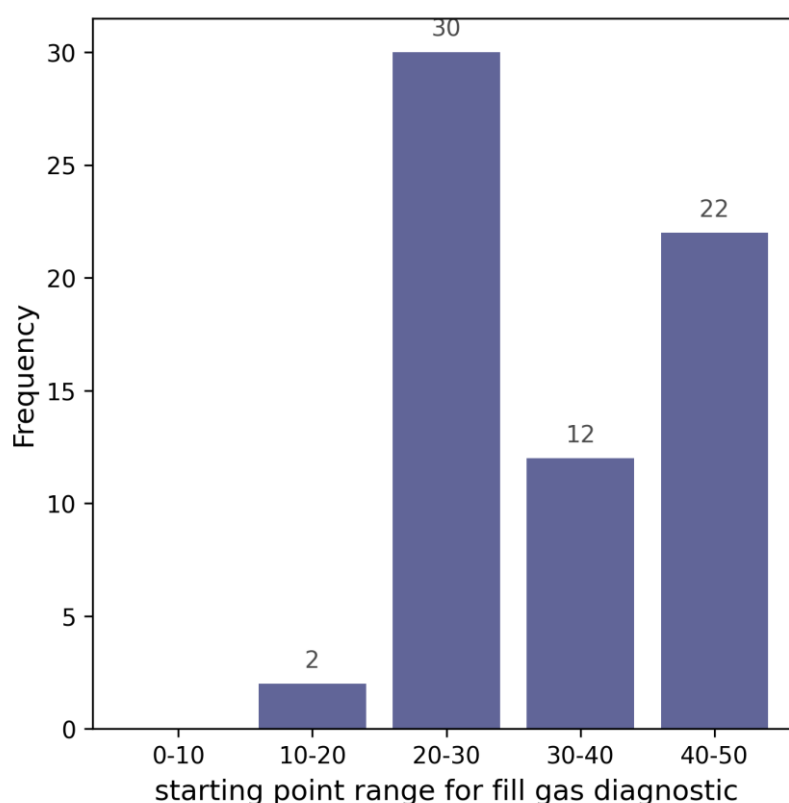
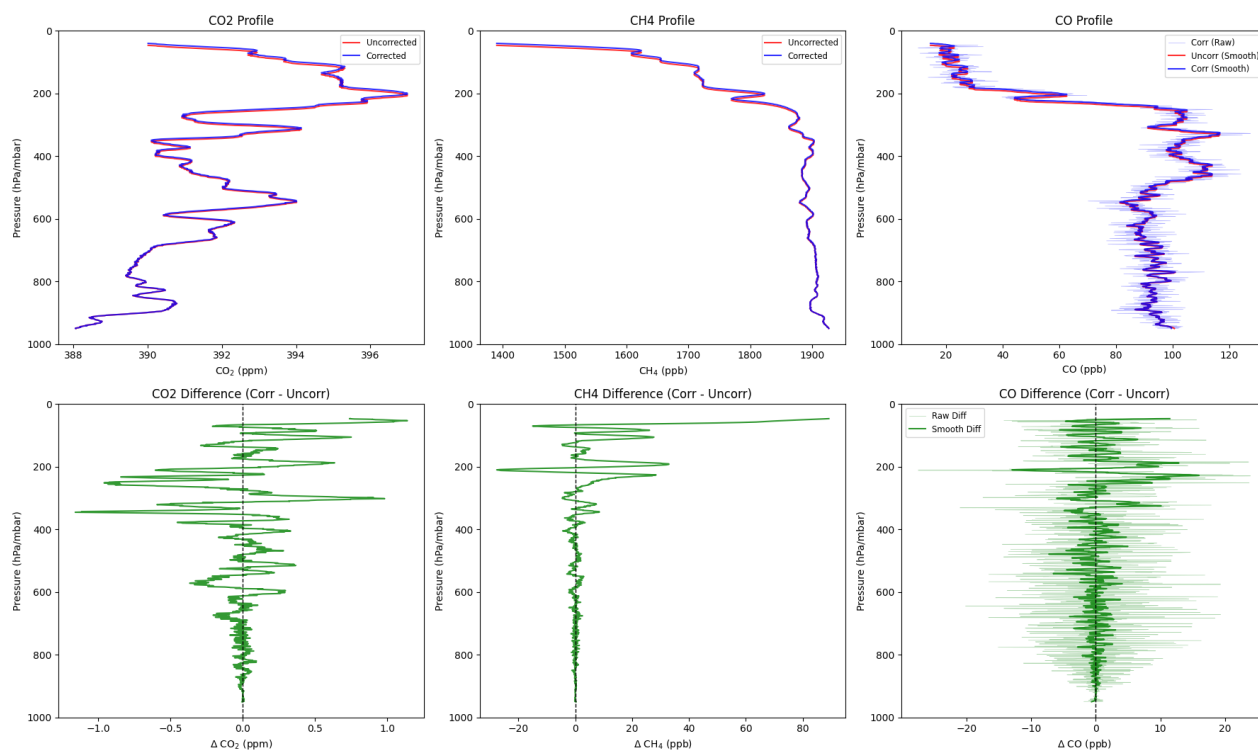


Figure 5. Distribution of diagnosed retrieval starting point for AirCore flights (2013-2025)

The distribution of diagnosed retrieval starting points for all AirCore flights during 2013–2025 is summarized in Fig. 5. Most flights exhibit optimized starting points between 20% and 50%, with the largest number of cases concentrated within the 20%–30% and 40%–50% ranges. The distribution demonstrates that the conventional assumption of a fixed 50% fill-gas transition point systematically overestimates the effective atmospheric sample starting point for a substantial fraction of AirCore flights. Instead, the results indicate that the optimal retrieval starting point varies between flights and is strongly influenced by the residual fill-gas fraction and the degree of pressure disequilibrium during sampling.

Overall, the statistical distribution confirms that the updated retrieval framework provides a more physically realistic description of the AirCore sampling process and improves the consistency of altitude registration across the long-term Sodankylä AirCore dataset.

AirCore Analysis: Sodankylä (2014-07-17)



AirCore Analysis: Sodankylä (2019-09-09)

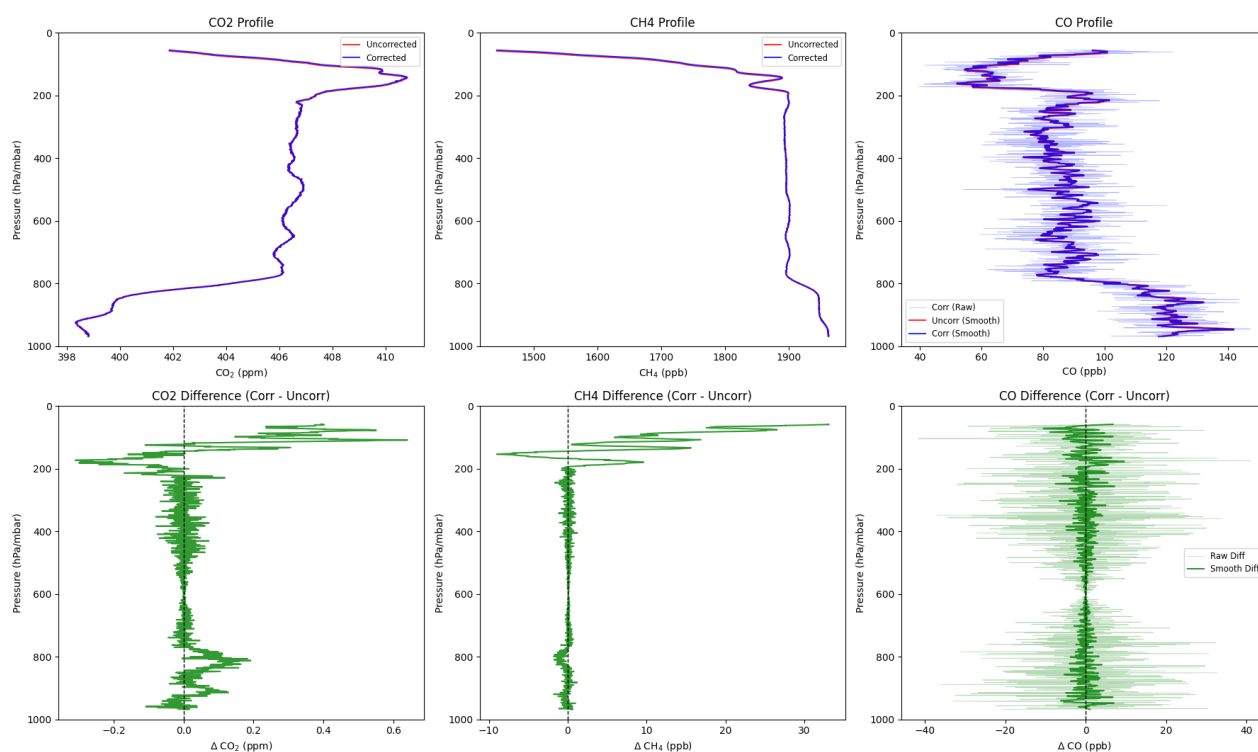


Figure 6. Flight comparison

Representative examples of the profile correction are shown in Fig. 6 for Flight 20140717 and Flight 20190909. The figure compares the original AirCore retrieval profiles with the updated retrieval profiles from Dataset V2.0 together with the corresponding LISA observations.

The updated retrieval procedure mainly affects the pressure-altitude registration of the upper tropospheric and stratospheric portions of the profiles. Compared with the original retrievals, the V2.0 profiles show improved vertical alignment with the independent LISA measurements, particularly for CH₄ and CO in regions with strong vertical gradients.

For the representative flights shown in Fig. 6, the revised retrieval framework reduces the altitude attribution bias and produces more physically consistent vertical structures. The correction generally shifts the atmospheric sample region toward lower starting-point percentages relative to the original 50% fill gas assumption, resulting in improved agreement with the observed tracer gradients.

Although the magnitude of the correction varies between individual flights, the retrieval improvements are consistently observed across the validation cases. The starting point corrected dataset has been used to generate NAF v2.0 (Figure 7).

 aircore_sod_20131022_5_v2.0.naf	2026/3/31 19:08	NAF 文件	294 KB
 aircore_sod_20140319_8_v2.0.naf	2026/3/31 14:17	NAF 文件	301 KB
 aircore_sod_20140409_9_v2.0.naf	2026/3/31 14:17	NAF 文件	299 KB
 aircore_sod_20140508_10_v2.0.naf	2026/3/31 14:17	NAF 文件	297 KB
 aircore_sod_20140714_11_v2.0.naf	2026/3/31 14:17	NAF 文件	298 KB
 aircore_sod_20140715_12_v2.0.naf	2026/3/31 14:17	NAF 文件	398 KB
 aircore_sod_20140716_13_v2.0.naf	2026/3/31 14:17	NAF 文件	367 KB
 aircore_sod_20140717_14_v2.0.naf	2026/3/31 14:17	NAF 文件	355 KB
 aircore_sod_20140718_15_v2.0.naf	2026/3/31 14:17	NAF 文件	342 KB
 aircore_sod_20150620_16_v2.0.naf	2026/3/31 14:41	NAF 文件	301 KB
 aircore_sod_20150623_19_v2.0.naf	2026/3/31 14:41	NAF 文件	299 KB
 aircore_sod_20160204_20_v2.0.naf	2026/3/31 19:15	NAF 文件	372 KB
 aircore_sod_20160310_21_v2.0.naf	2026/3/31 19:15	NAF 文件	364 KB
 aircore_sod_20160414_22_v2.0.naf	2026/3/31 19:15	NAF 文件	362 KB
 aircore_sod_20160712_24_v2.0.naf	2026/3/31 19:15	NAF 文件	352 KB
 aircore_sod_20160824_25_v2.0.naf	2026/3/31 19:15	NAF 文件	365 KB
 aircore_sod_20170421_27_v2.0.naf	2026/3/31 19:25	NAF 文件	301 KB
 aircore_sod_20170424_28_v2.0.naf	2026/3/31 19:25	NAF 文件	299 KB
 aircore_sod_20170515_30_v2.0.naf	2026/3/31 19:25	NAF 文件	297 KB

Figure 7. AirCore dataset V2.0

5.2 Description of AirCore retrieval procedure

The AirCore retrieval process reconstructs atmospheric vertical profiles by mapping the measured trace gas mole fraction time series to atmospheric pressure and altitude coordinates. The retrieval framework combines AirCore housekeeping measurements, meteorological observations, and high-frequency trace gas measurements to determine the temporal and spatial relationship between sampled air parcels and their corresponding atmospheric locations. The updated AirCore retrieval procedure used in Dataset V2.0 was specifically developed to improve altitude attribution accuracy, particularly in the stratosphere where pressure disequilibrium during sampling can introduce significant biases in the reconstructed profiles.

The AirCore retrieval process maps trace gas mole fractions to sampling altitudes to reconstruct vertical profiles. It consists of four steps: **(1) PREPROCESS**: Aligns AirCore housekeeping data (e.g., pressure, temperature, GPS position) with ambient meteorological measurements and selects the descent period for further analysis. Calibration of raw sensor signals (e.g., pressure, temperature) is also performed at this stage. **(2) QA/QC**: Curates the raw and processed files, applies trace gas calibration, identifies valid measurement periods, and removes contaminated measurements at the top and bottom of the profiles to ensure the quality of the final data product. **(3) RETRIEVAL**: Matches the sampling and analysis time series to derive vertical profiles of trace gases based on a pressure–altitude mapping, using a flow model and assuming ideal gas behavior. **(4) SENSITIVITY TESTS**: Optimize the AirCore retrievals by adjusting the starting point of the mole fraction time series to reduce altitude registration uncertainties.

Compared with the previous retrieval framework, the updated procedure introduces two major modifications related to altitude registration. First, during the PREPROCESS stage, a new diagnostic method is implemented to estimate the optimal retrieval starting point based on the relative duration of the residual fill gas interval within the total AirCore analysis period. Second, during the RETRIEVAL stage, the atmospheric sample profile is reconstructed beginning from the measured balloon top pressure instead of the conventional 0 hPa reference used in earlier retrieval approaches.

5.2.1 PREPROCESS

The PREPROCESS stage prepares all required datasets for profile reconstruction and determines the temporal structure of the AirCore sampling sequence. The system first imports and synchronizes AirCore housekeeping data, including pressure, temperature, GPS position, and operational flight status, together with meteorological observations obtained during the balloon flight. Flight phases are classified according to the meteorological pressure record, with separate identification of ascent and descent periods. The descent portion of the flight is selected for retrieval because this period corresponds to the atmospheric sampling process inside the AirCore tube.

The preprocessing stage also includes the calibration and correction of the raw trace gas measurements obtained from the Picarro cavity ring-down spectroscopy (CRDS) analyzer (Fig 8, Fig 9). Calibration gas measurements before and after the AirCore analysis are used to apply one-point corrections for CO₂ and CH₄ and linear corrections for CO. During this step, the AirCore analysis interval is identified, including the start and end times of the complete AirCore measurement sequence.

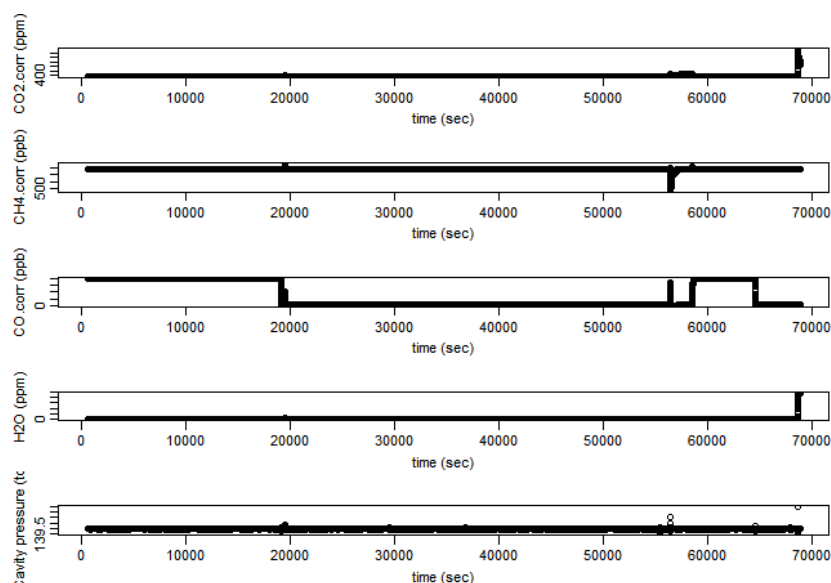


Figure 8. Raw CO₂, CH₄, CO, H₂O and Cavity Pressure time series of AirCore samples through Picarro CRDS Analyzer.

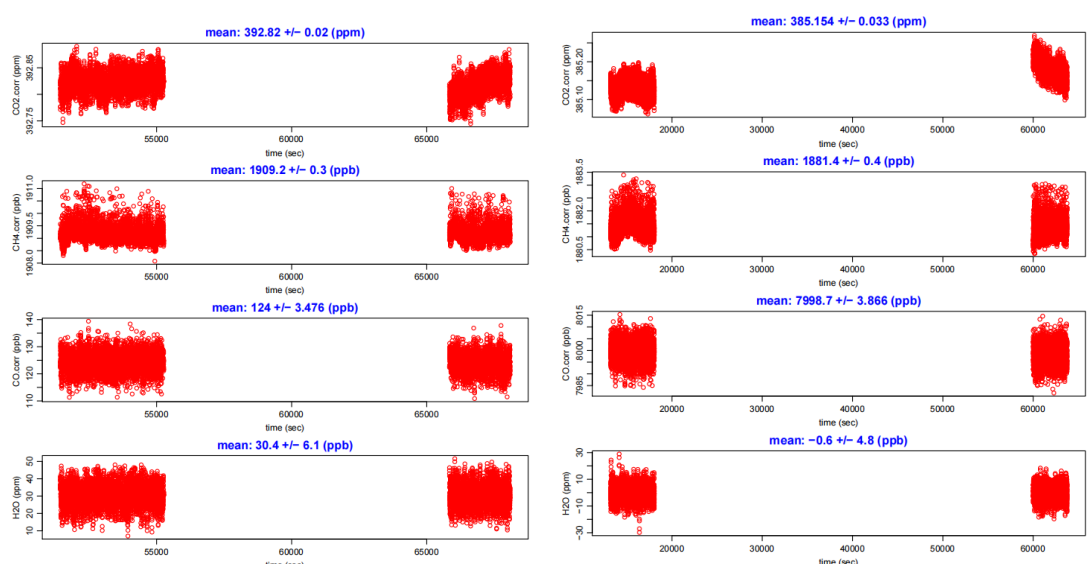


Figure 9. Manually selected information on the temporal position and mole fractions of fill gas (right) and calibration gas (left) in the Picarro CRDS trace gas time series. This information is used to correct for potential nonlinearity in the instrument's response or discrepancies in calibration over time.

The AirCore analysis sequence consists of several distinct gas components measured consecutively by the analyzer: push gas (PG), calibration gas (Cal gas), residual fill gas (FG), atmospheric sample air (Fig 9).

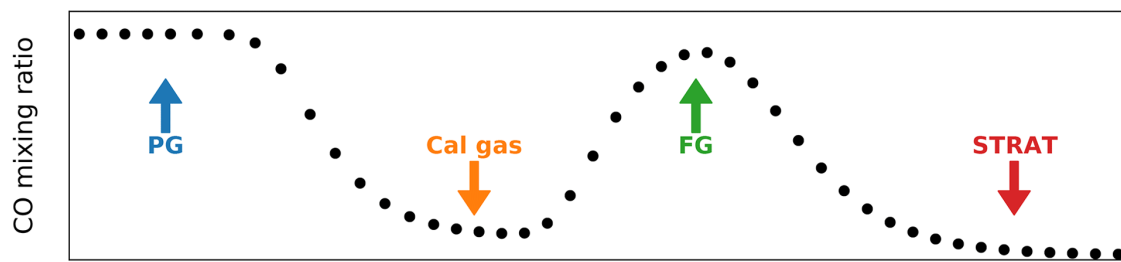


Figure 9. AirCore analysis sequence

The fill gas region is particularly important because it defines the transition between residual gas remaining inside the AirCore tube and the actual atmospheric sample. In the CO time series, this transition produces a characteristic high-CO peak that has traditionally been used to define the retrieval starting point.

In earlier retrieval algorithms, the sample starting point was defined using the half-maximum location of the fill gas peak, corresponding to the 50% transition point between fill gas and atmospheric sample air. This approach implicitly assumed instantaneous pressure equilibrium between the AirCore tube and ambient atmosphere during sampling.

However, the updated retrieval framework recognizes that residual fill gas does not actively participate in the atmospheric gas exchange process during balloon descent. Therefore, the residual fill gas transition should remain fixed, while the effective atmospheric sample starting point may vary between flights depending on sampling dynamics and pressure disequilibrium effects.

To account for this behavior, the PREPROCESS stage introduces a new diagnostic starting-point estimation method based on the relative duration of the fill gas interval. The duration of the residual fill gas period is calculated from the analyzer time series using the identified sample starting index (`sample.index`) and the AirCore analysis interval (`aircore.index`). The ratio between fill gas duration and total AirCore analysis duration is then used together with the measured surface pressure and balloon top pressure to estimate the effective retrieval starting point for each flight.

The updated algorithm determines the sample starting time (`time.fill.end`) using a variable fill gas fraction rather than a fixed 50% threshold:

```
time.fill.end<-approx(dat[sel,"co.corr"],dat[sel,"time"],comax*1*0.01,rule=1)$y
```

where `1` represents the optimized starting-point percentage diagnosed from the fill gas analysis.

To facilitate later retrieval diagnostics, additional profile flags are introduced into the processed dataset, including:

- `aircore.index`: complete AirCore analysis interval,
- `sample.index`: atmospheric sample interval,
- `top.index`: top-of-profile region,
- `bottom.index`: bottom-of-profile region.

These indices are stored in the final processed files and allow direct calculation of fill gas duration, atmospheric sample duration, and retrieval boundaries for each flight. This step will determine the start and end times of the AirCore measurement in the analyzer time series (Fig 10, Fig 11) and

generate the picarro aircore data file (picarro_YYYYMMDD_FlightNO_aircore.csv, picarro_YYYYMMDD_FlightNO_calibrated.csv).

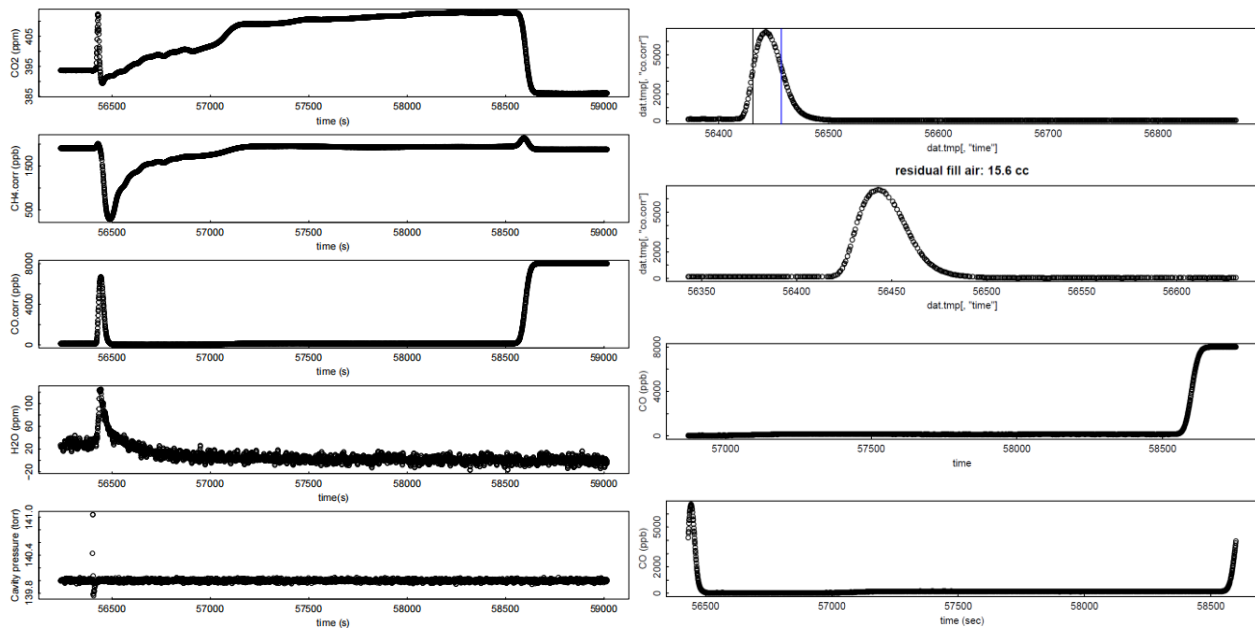


Figure 10. The mixing part of CO fill gas and stratospheric sample, the blue line represents the zoom version of the starting point. This point can calculate the residual fill gas volume in the tube. This process still defines the midpoint of the fill gas as the endpoint, shown in the second black line.

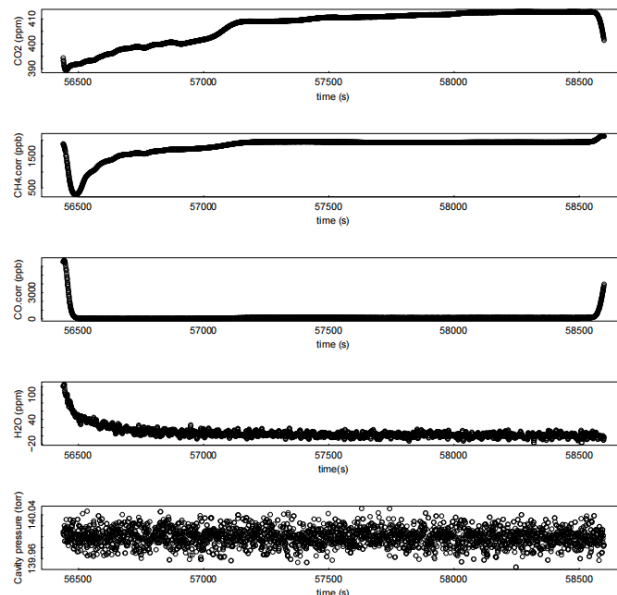


Figure 11. Time series of sample CO₂, CH₄, and CO mole fractions on picarro time series. Trace gas data aligned with corrected altitudes for further retrieval processes.

During meteorological preprocessing, the updated retrieval framework additionally determines: surface pressure (p.ground), balloon top pressure (p.top), from the descent portion of the flight:

```
p.ground<-max(dat.met[, "P.amb"]); p.top <- min(dat.met[, "P.amb"]);
```

The diagnosed balloon top pressure becomes a key parameter in the revised retrieval framework. This step will generate the Meteorological Data File containing the meteorological and GPS data (metdata_YYYYMMDD_FlightNO.csv). This file contains all original and smoothed meteorological and GPS metadata like temperature, humidity, altitude, etc. And GPS data like velocity, latitude, longitude, etc. It also contains up to six columns for temperature sensor measurements (T.coil.1-6) from the AirCore. The smoothed version of the Temperature value is the average (T.coil) of the 6. "flight" indicates whether the AirCore is ascending (1), descending (2), or on the ground (0).

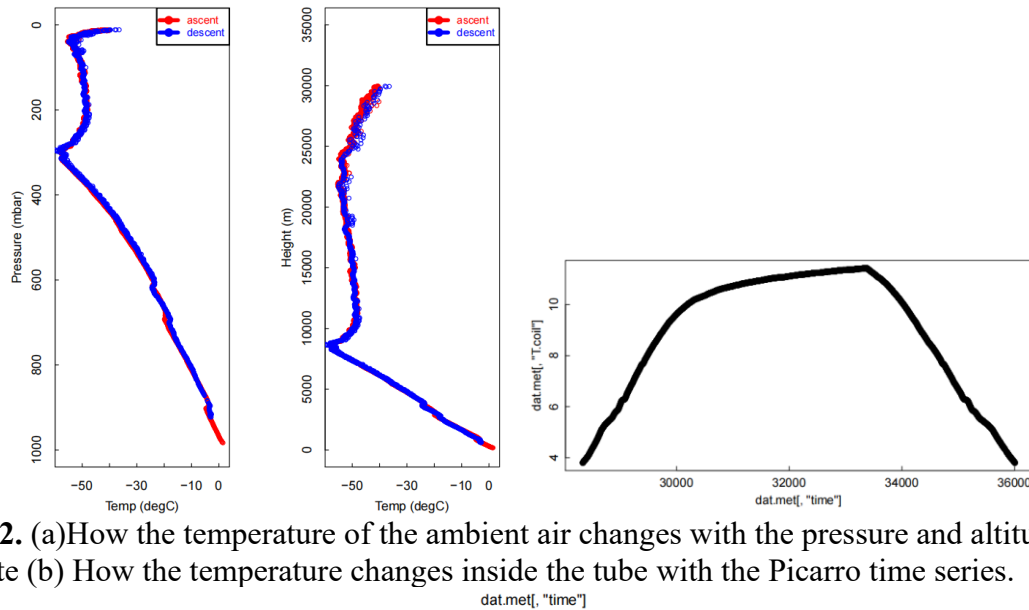


Figure 12. (a) How the temperature of the ambient air changes with the pressure and altitude coordinate (b) How the temperature changes inside the tube with the Picarro time series.

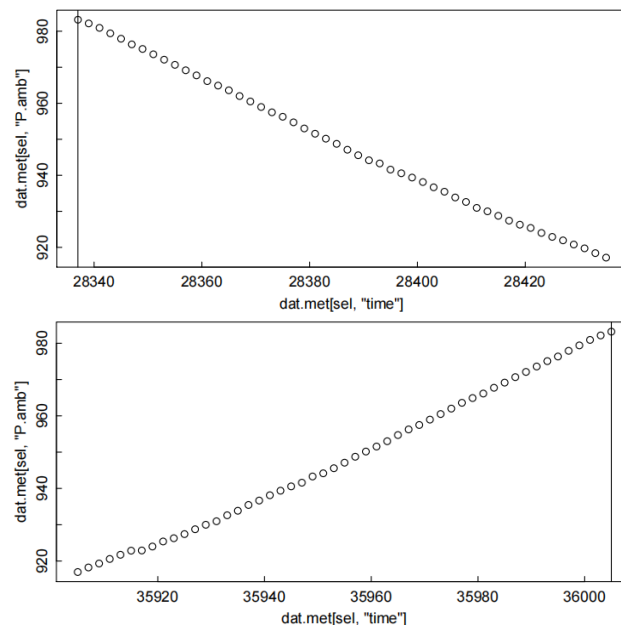


Figure 13. How the temperature of the ambient air changes as AirCore takes off and lands, and the vertical lines indicate the points in time when AirCore begins to take off and land on the Picarro time series.

5.2.2 QA/QC

The quality assurance and quality control (QA/QC) stage ensures the consistency and reliability of the retrieved AirCore profiles. During this stage, calibrated trace gas measurements are screened for instrumental artifacts, outliers, and contamination events. The valid AirCore analysis interval is verified using the characteristic gas sequence observed in the CO time series.

Contaminated measurements at both the top and bottom portions of the AirCore profiles are identified and removed. These contaminated regions are typically associated with mixing between atmospheric sample air and calibration or purge gases during the beginning and end of the analysis process.

The QA/QC stage also verifies the consistency of the flight metadata, retrieval boundaries, and time synchronization between meteorological and analyzer datasets. The resulting processed files contain calibrated trace gas mole fractions together with retrieval metadata and profile classification indices.

5.2.3 RETRIEVAL

The RETRIEVAL stage converts the time-resolved trace gas measurements into pressure-resolved atmospheric vertical profiles. In the traditional AirCore retrieval framework, profile reconstruction was performed by assuming that the atmospheric sample began at the midpoint of the fill gas transition region and that the pressure profile extended continuously from 0 hPa to surface pressure.

In the updated retrieval procedure V2.0, the atmospheric sample retrieval instead begins from the measured balloon top pressure (p.top) obtained from the AirCore data logger and meteorological observations. This represents one of the major conceptual differences from the previous retrieval framework.

The updated retrieval first reconstructs the residual fill gas region separately between 0 hPa and the balloon top pressure. The pressure distribution within the fill gas segment is generated using the relative analyzer time fractions:

```
pressure<-0:p.ground; p.change <- max(pressure)- min(pressure);
```

The atmospheric sample region is then reconstructed beginning from the balloon top pressure and extending downward toward surface pressure:

```
pressure<-p.top:p.ground; p.change <- max(pressure)- min(pressure);
```

This modification reflects the physical interpretation that atmospheric sampling only begins after the residual fill gas region has been displaced during descent.

The retrieval algorithm uses the temporal evolution of analyzer measurements together with the pressure history of the balloon descent to assign atmospheric pressure coordinates to each measured air parcel. Because equal pressure changes do not correspond to equal sampled air masses under varying temperatures, temperature corrections are included using the ideal gas law.

During descent, the amount of sampled air associated with each pressure level depends on both pressure and AirCore tube temperature. Up to six internal coil temperature sensors are averaged to generate a representative tube temperature profile (T.coil). These temperatures are interpolated onto the reconstructed pressure coordinates and used to calculate the corresponding molar air content inside the AirCore tube.

The retrieval therefore accounts for thermal expansion effects during sampling and produces temperature-corrected pressure coordinates (p.prof.coil.T.corr) for the final atmospheric profiles.

After reconstruction of both the residual fill gas region and atmospheric sample region, the two profile segments are merged to generate the final continuous AirCore vertical profile. The resulting output files contain: reconstructed pressure coordinates, temperature-corrected pressure profiles, calibrated CO₂, CH₄, and CO mole fractions, profile classification indices, timing metadata.

The updated retrieval products are written as individual flight files (aircore_YYYYMMDD_FlightNO_corrected).

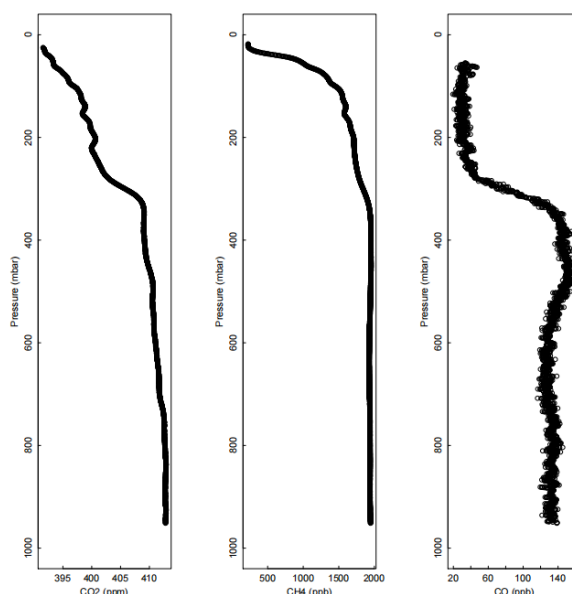


Figure 14. profiles of AirCore CO₂, CH₄, and CO in pressure coordinate

5.2.4 SENSITIVITY TESTS

The final stage of the retrieval framework consists of sensitivity tests designed to quantify uncertainties associated with altitude attribution and starting-point determination.

For the six flights with simultaneous AirCore and LISA observations, multiple retrievals are generated using different starting-point percentages. Each retrieved profile is compared with the corresponding LISA measurements, and weighted RMSE statistics are calculated to determine the optimal starting-point range.

5.2.4.1 Deconvoluted method to extract AirCore peak signals

AirCore measurements undergo molecular diffusion as they are collected, during storage and analysis, resulting in smoothed concentration profiles. The signal retrieved from AirCore is not the original high-resolution profile but a smoothed version due to diffusion effects during the sampling, storage, and analysis process. From the measurements of AirCore and LI-COR in Figure 15, it can be seen that this smoothing dilutes peak values and alters the sharpness of the original gas concentration signals, complicating direct comparison with other datasets such as LISA. Figure 16 shows deconvolution can reconstruct most peak signals of AirCore retrieval profiles and correct for mixing and smoothing effects.

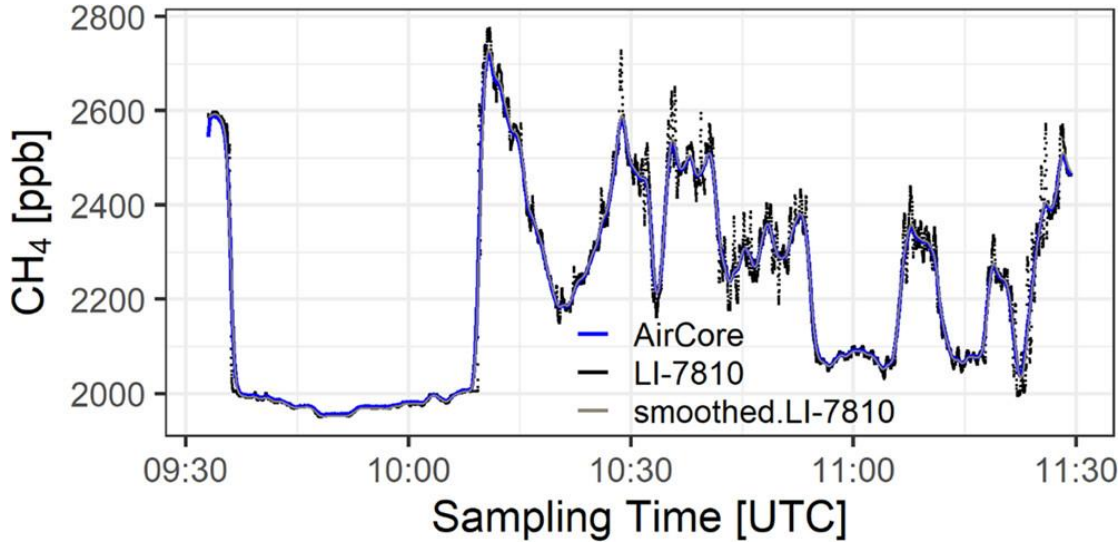


Figure 15. Comparison of AirCore and LI-7810 measurements for CH₄ from Tong et al.(2023) To retrieve the original, high-resolution AirCore profile from the smoothed data, a deconvolution method is applied. The deconvolution process works by applying a model that accounts for the diffusion that occurs during AirCore sampling. The formula is as follows:

$$c(t) = (g * s)(t) = \int_0^t \frac{1}{\tau} e^{-(t-t')/\tau} s(t') dt'$$

Where $g(t) = \frac{1}{\tau} e^{-t/\tau}, t \geq 0$

t is the current observation time, t' is a past time instant traversed during integration, $t - t'$ is the time difference between the current time and the past time instant.

$$s(t) = c(t) + \tau \frac{dc(t)}{dt}$$

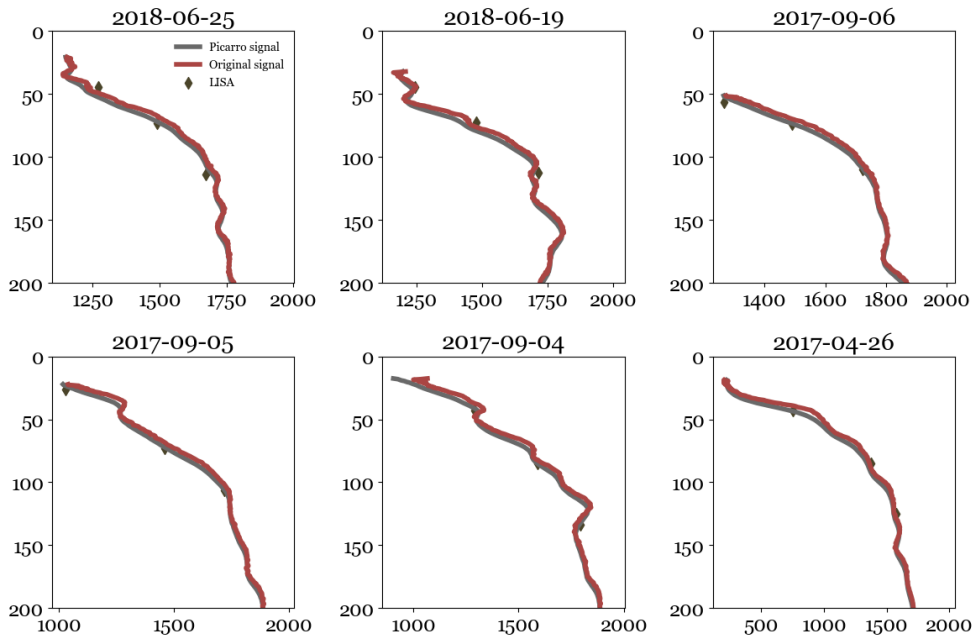


Figure 16. Results for CH₄ profiles with (red line) and without (grey line) deconvolution.

5.2.4.2 LISA constrained method

The sensitivity tests additionally compared two different retrieval frameworks for pressure initialization:

Scheme A: retrieval beginning from 0 hPa, in which both the residual fill gas reference point and the atmospheric sample starting point are shifted simultaneously;

Scheme B: retrieval beginning from the measured balloon top pressure (p.top), in which the residual fill gas reference remains fixed and only the atmospheric sample starting point is adjusted.

For each of the six AirCore–LISA comparison flights, a series of retrievals was generated using different starting-point percentages, and the corresponding weighted RMSE values were calculated relative to the LISA observations. The resulting “starting point versus RMSE” relationships are shown in Fig. 17.

The comparison reveals that Scheme A frequently produces lower RMSE values when the retrieval starting point exceeds approximately 40%. However, this behavior does not necessarily indicate that retrievals initialized from 0 hPa provide more accurate altitude attribution. Instead, the reduced RMSE at large starting points is interpreted as a coordinate compensation effect rather than a physically improved retrieval solution.

In Scheme A, the entire pressure coordinate system shifts together with the retrieval starting point because the residual fill gas reference point is allowed to move simultaneously with the atmospheric sample region. As a result, part of the altitude registration error introduced at large starting-point offsets is artificially compensated by the displacement of the fill gas zero-pressure reference. Under these conditions, the expected increase in RMSE caused by pressure-altitude mismatch is partially masked.

By contrast, Scheme B preserves a fixed residual fill gas reference and initializes the atmospheric profile from the independently measured balloon top pressure. In this framework, changes in the retrieval starting point directly affect the atmospheric pressure mapping without modifying the reference coordinate system itself. Consequently, the RMSE response more clearly reflects the true altitude attribution uncertainty associated with the retrieval.

The results demonstrate that Scheme A is relatively insensitive to starting-point variations because the retrieval effectively modifies the reference coordinate system rather than strictly preserving physical consistency in the pressure-altitude relationship. Allowing the fill gas reference point to move freely tends to flatten the RMSE response curve and can therefore obscure the true magnitude of altitude registration errors.

In contrast, Scheme B consistently produces a distinct RMSE minimum for all six validation flights, with optimized starting points concentrated within a relatively narrow range. The existence of a clear and physically interpretable RMSE minimum indicates that the retrieval framework initialized from balloon top pressure provides a more robust and physically consistent representation of the AirCore sampling process. Therefore, Scheme B was selected as the reference retrieval framework for AirCore Dataset V2.0.

Figure 18 compares Scheme B with Scheme D, which applies deconvolution before altitude attribution and then follows the Scheme B retrieval procedure. While Scheme B adjusts the pressure

mapping, Scheme D aims to correct instrumental signal smoothing. The optimized start point for Scheme D is $46\% \pm 9\%$. Compared with Scheme B, deconvolution reduces the RMSE for only two flights (2017-04-26 and 2018-06-19), while increasing the RMSE for the other four flights. This indicates that deconvolution does not consistently improve retrieval accuracy. Therefore, Scheme B was retained as the reference framework for AirCore Dataset V2.0, with deconvolution treated as an additional correction rather than a universally beneficial step.

Scheme C aims to correct pressure biases in the Scheme A retrieval using an empirical relationship between the pressure difference inside and outside the AirCore tube and the ambient pressure. Preliminary simulations using the Tans (2022) AirCore filling dynamics model suggests a relationship between pressure bias and ambient pressure, but the validity of the simulated relationship requires further verification. Based on the simulated pressure bias, a linear correction was constructed over the 0–200 hPa range, with zero bias assumed at 200 hPa. For each candidate retrieval start point shift, the pressure difference at the shifted profile top was used to define the upper endpoint of the linear correction, while the 200hPa reference provided the lower endpoint. The correction was applied only within the 0–200 hPa range, and the result profiles were compared with those from Scheme B, which adjusts the retrieval start point without applying this empirical pressure correction. Figure 19 shows the comparison between Scheme B and C based on the simulated pressure differences.

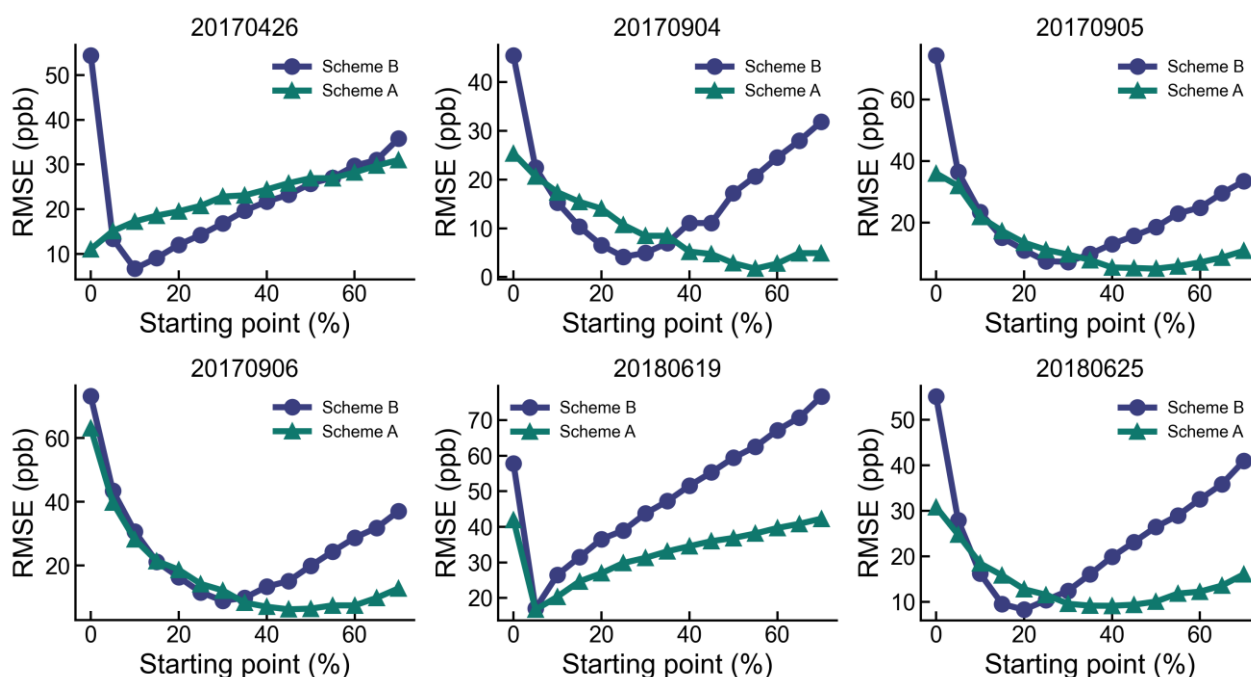


Figure 17. Two retrieval frameworks Scheme A vs. Scheme B for six flights in Sodankylä

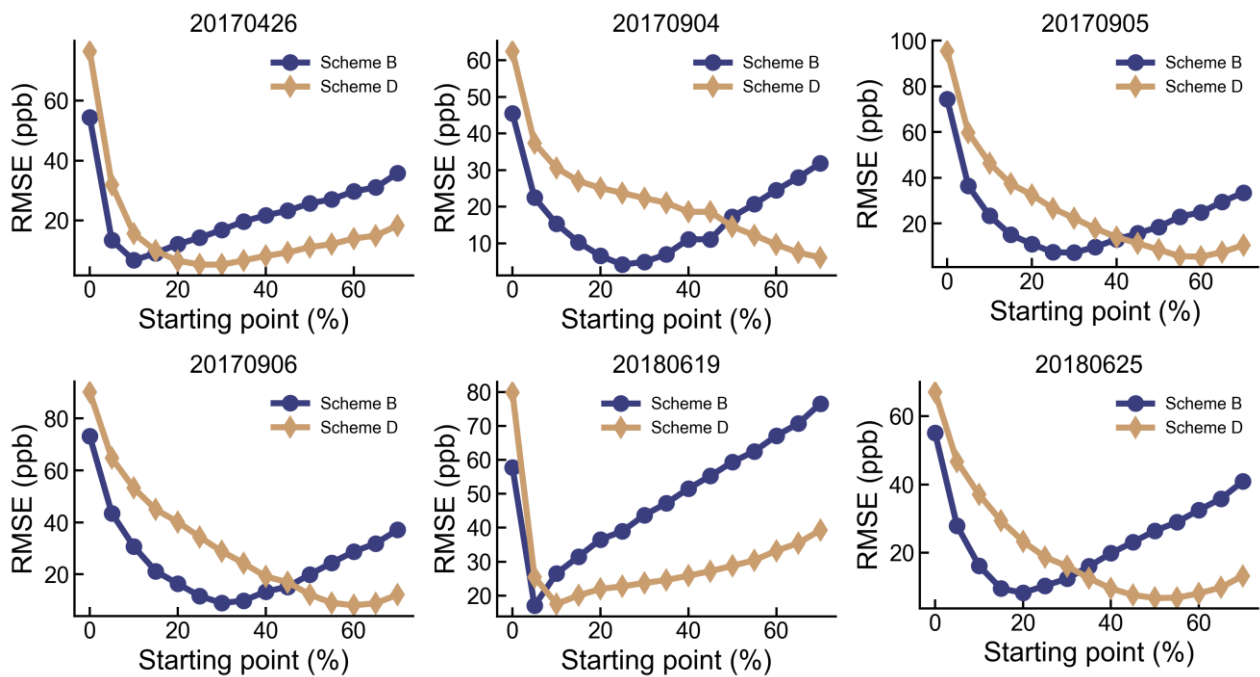


Figure 18. deconvolution vs. starting point optimization

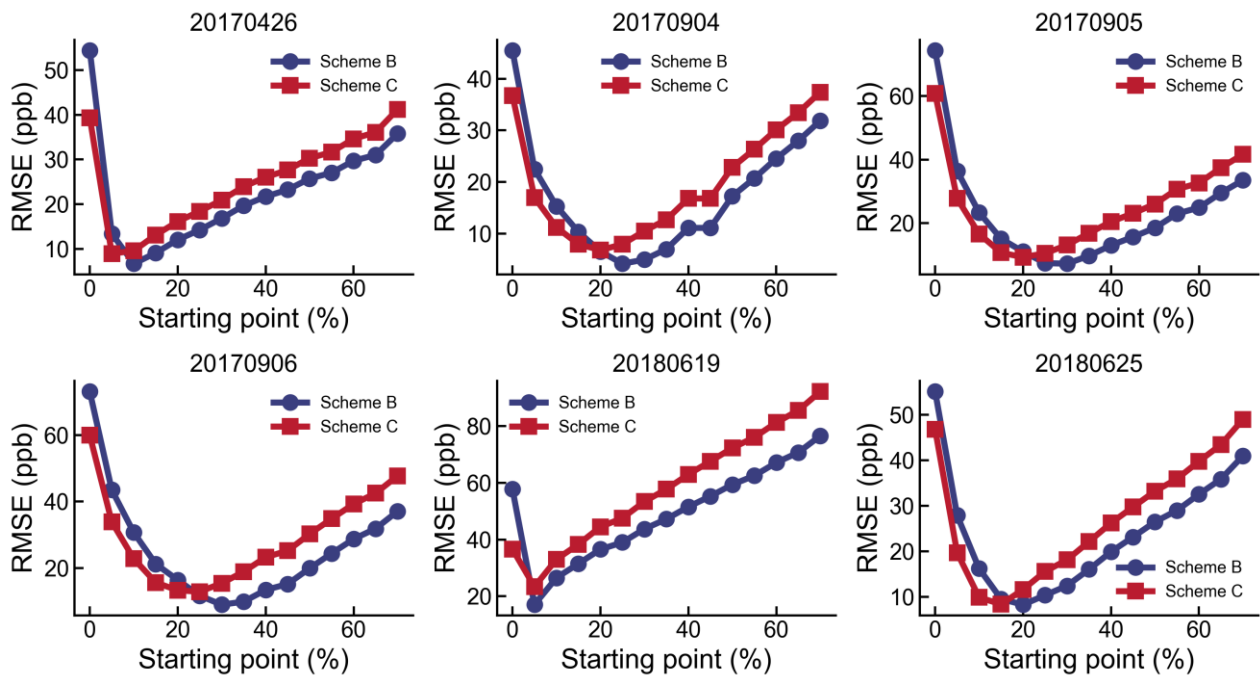


Figure 19. The optimization results for comparison between AirCore and LISA

5.2.4.3 Tans' pressure correction method

Developed by Dr. Pieter Tans, this method models the fill process dynamics to associate air parcel positions with sampling altitudes.

Requires accurate modelling of the pressure drop across valves and desiccators during AirCore descent. Includes parameters like C_v (flow coefficient) and XTPR (terminal pressure drop ratio) to represent the valve fluid dynamics.

Profiles of CO₂, CH₄, and CO against pressure coordinates for June 19, 2018, have been drawn using Tan's pressure correction method and will be compared to the original AirCore profiles as well as to LISA in the next step.

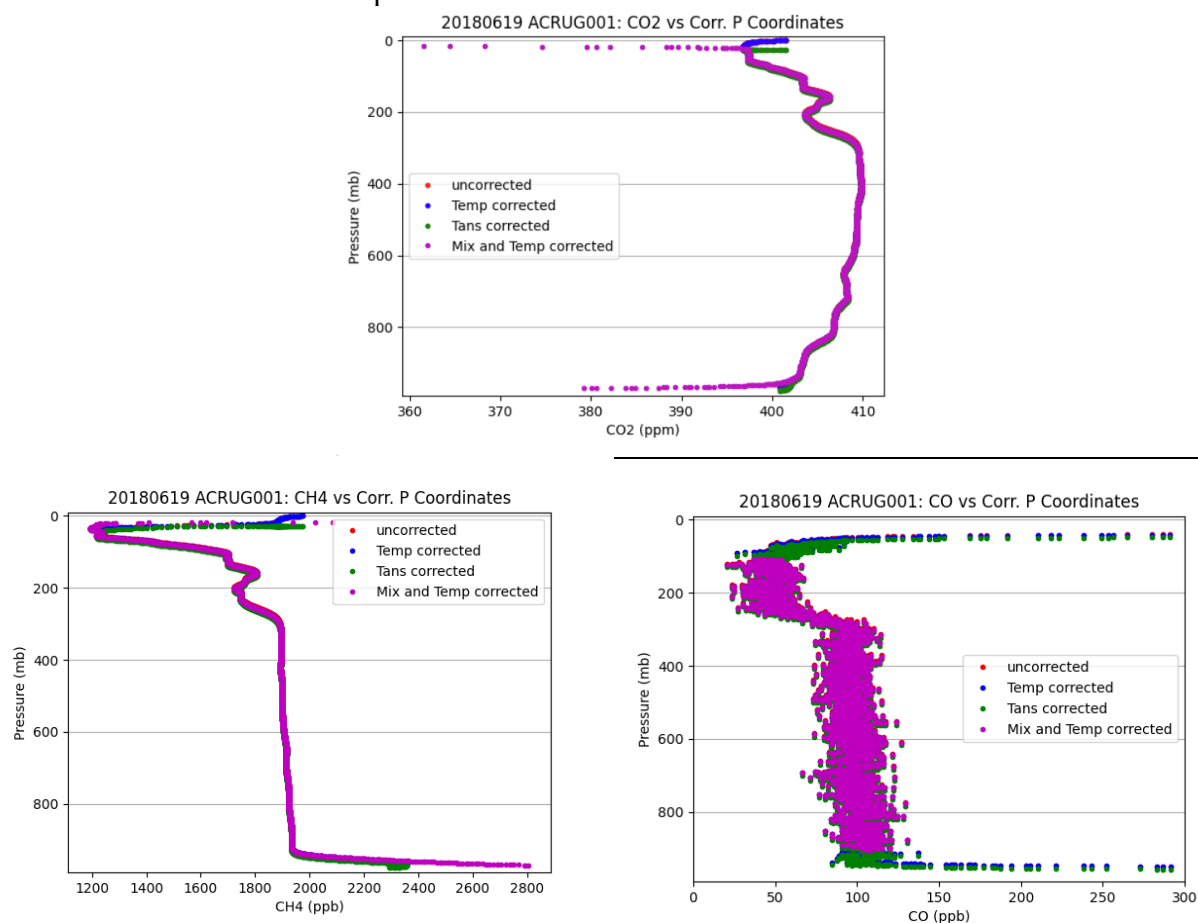


Figure 20. Results for CO₂, CH₄, CO against corrected pressure coordinates

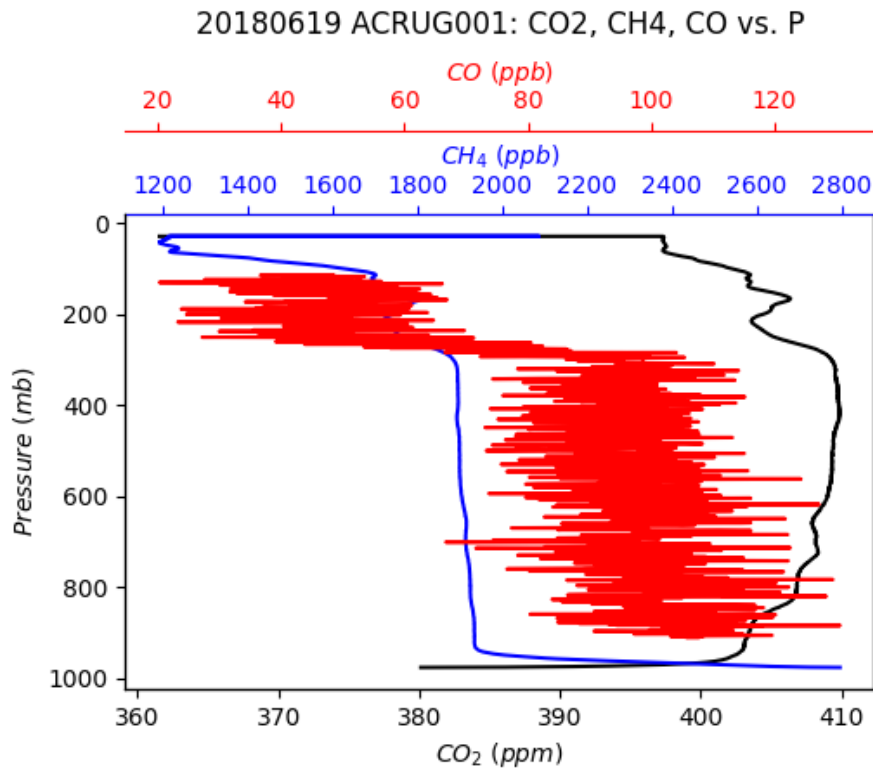


Figure 21. Plot CO₂, CH₄, CO against pressure height

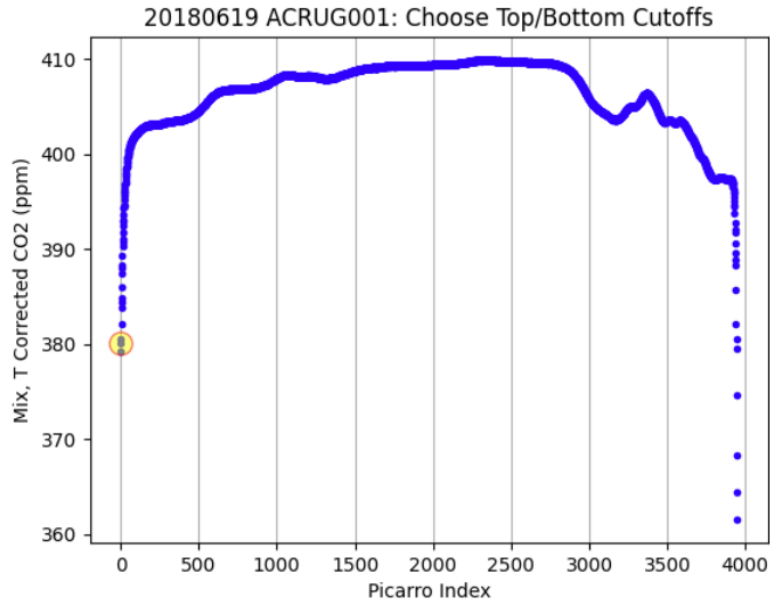


Figure 22. Select AirCore Top and Bottom Index

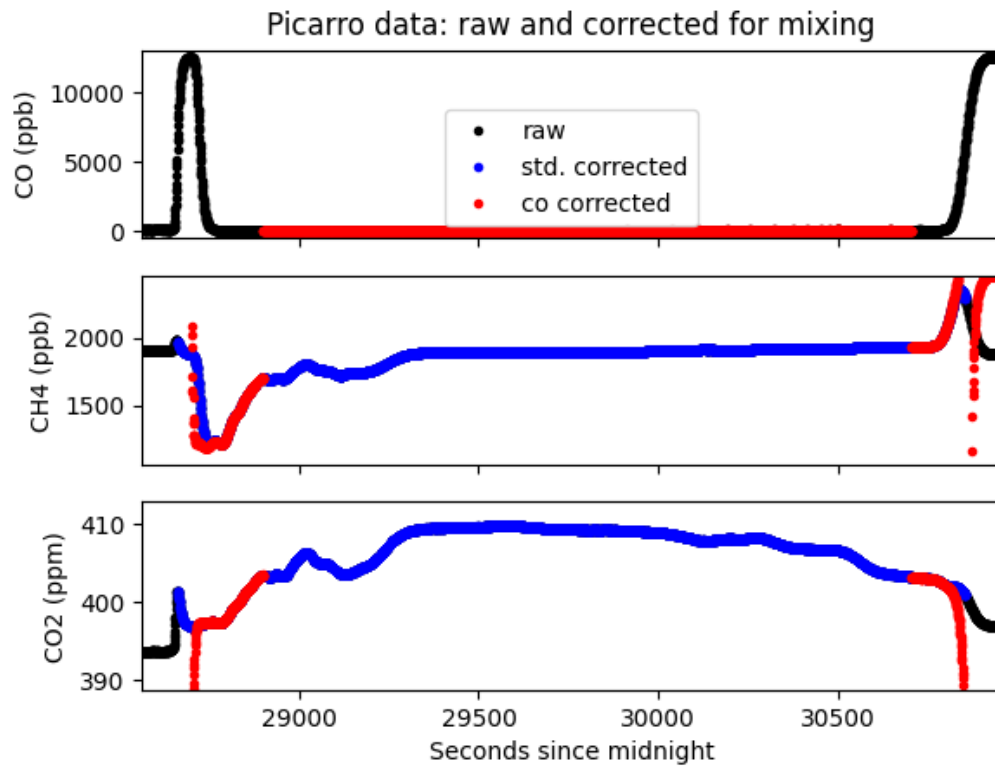


Figure 23. Result for mixing corrected CO₂, CH₄ and CO

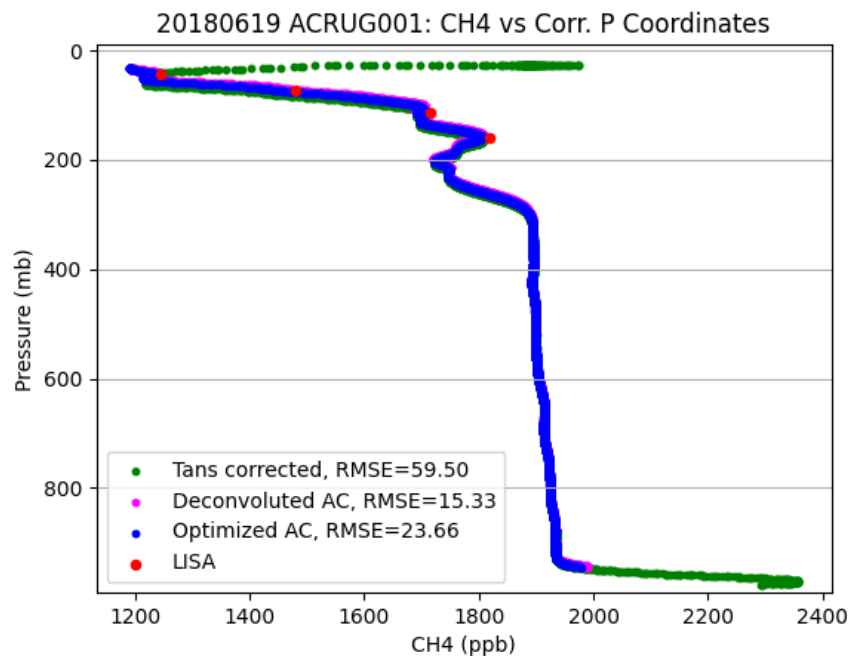


Figure 24. Comparison of CH₄ Profiles: Tans Correction, Optimized AirCore, and Deconvolution vs. LISA Sample

6 OUTLOOK

The updated AirCore Dataset V2.0 will be made available and used in scientific studies, such as validating remote sensing measurements and investigating stratospheric circulations. Feedback will be collected and used to further improve the AirCore dataset V2.0.

7 APPLICABLE DOCUMENTS

Statement of Work: Fiducial Reference Measurements for Ground-Based IR Greenhouse Gas Observations (FRM4GHG 2.0)
Prepared by: EOP-GMQ, Reference: ESA-EOPG-EOPGMQ-SOW-21

8 REFERENCE DOCUMENTS

Tong, X., van Heuven, S., Scheeren, B., Kers, B., Hutjes, R., & Chen, H: Aircraft-Based AirCore Sampling for Estimates of N₂O and CH₄ Emissions. *Environmental science & technology*, 57(41), 15571-15579, 2023

Tans, P.: Fill dynamics and sample mixing in the AirCore, *Atmos. Meas. Tech.*, 15, 1903–1916, <https://doi.org/10.5194/amt-15-1903-2022>, 2022.

Wagenhäuser, T., Engel, A., and Sitals, R.: Testing the altitude attribution and vertical resolution of AirCore measurements with a new spiking method, *Atmos. Meas. Tech.*, 14, 3923–3934, <https://doi.org/10.5194/amt-14-3923-2021>, 2021.