

FRM4GHG 2.0

Fiducial Reference Measurements for Greenhouse Gases



Technical Note

LHR data retrieval

Deliverable: D2.2
Date: 01 Oct 2024
Lead author(s): D. Weidmann, Matt Berrington, Neil Macleod, A. Dudhia
Subject: ESA-Contract No. 4000136108/21/I-DT-lr
Category: ESA Express Procurement (EXPRO)
Our ref.: RS0

Table of contents

1	Document change record	3
2	Access list.....	3
3	Purpose & background	3
4	Document structure.....	3
5	Level 1 data processor.....	3
5.1	Overall architecture	3
5.2	Functional blocks description	4
5.2.1	Level 0 data reading	4
5.2.2	Sun geometric elevation	5
5.2.3	Relative wavenumber calibration	5
5.2.4	Background corrections	6
5.2.5	Non-linearity correction	8
5.2.6	Absolute wavenumber calibration.....	8
5.2.7	Quality flagging	10
5.2.8	L1 data file	10
5.3	L1 data improvement summary	11
6	Level 2 processor	13
6.1	L1C pre-processor	13
6.2	Multispectral Orbital Retrieval using Sequential Estimation.....	13
6.3	Inputs.....	14
6.4	Preliminary example run	14
7	Conclusion and forward look	15
7.1	Full processor update.....	15
7.2	Future instrumental and operational changes.....	16
1	Applicable documents	16
2	Reference documents	17

1 DOCUMENT CHANGE RECORD

Issue	Date	Item	Comment
1	14/10/2024	All	First release within FRM4GHG

2 ACCESS LIST

This document is a deliverable “D2.2 TN on LHR data retrieval” – 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 & BACKGROUND

The document describes the modifications made to the Laser Heterodyne spectro-Radiometer (LHR) data processors allowing for the retrieval of gas volume mixing ratios from the instrument raw output data.

Since FRM4GHG 1, the decision was made to fully overhaul the retrieval software in order to:

- Account for the instrument modification and specifically the inclusion of a new VIS/IR hyperspectral imager (HSDI for Hyperspectral Solar Disk Imager).
- Harmonize the retrieval software in line with the approach implemented for a limb solar occultation mission (ESA SCOUT CubeMAP, now UKSA SOLSTICE).

4 DOCUMENT STRUCTURE

See Table of Contents.

5 LEVEL 1 DATA PROCESSOR

5.1 Overall architecture

The L1 processor is written in the IDL language.

Throughout the processor NetCDF format is used. The datafile granularity is one day, and one datafile per instrument.

The processor has been written for compatibility with a L2 processor based on the Reference Forward Model – RFM (Dudhia, 2017).

It takes three files as inputs:

- The Laser Heterodyne spectro-Radiometer (LHR) L0 data file. Note that the LHR is sometimes referred as HIROS (High-resolution IR Occultation Spectrometer), owing to the heritage from the CubeMAP mission.
- The Hyperspectral Solar Disk Imager (HSDI) L0 data file.

- A configuration file.

The flow block diagram of the L1 processor is given in Fig. 5.1.

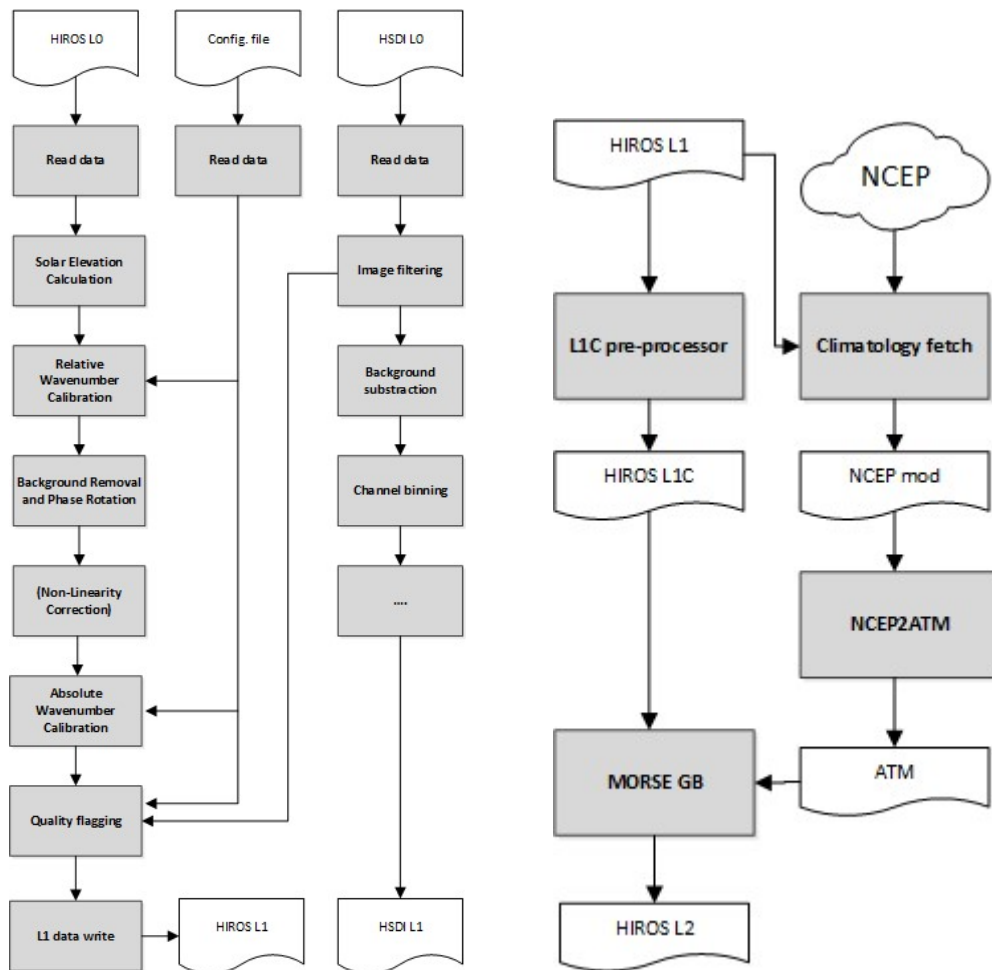


Fig. 5.1: Functional block diagram of the Level 1 (left) and level 2 (right) data processor

5.2 Functional blocks description

5.2.1 Level 0 data reading

The L0 NetCDF files, with a single level grouping shown in Fig. 5.2, are read with standard NetCDF libraries. They are then organized into structured data (IDL structure format) ready to be utilized by the rest of the processor. A typical LHR L0 file is up to 50 MB. A typical HSDI L0 file is up to 4 GB.

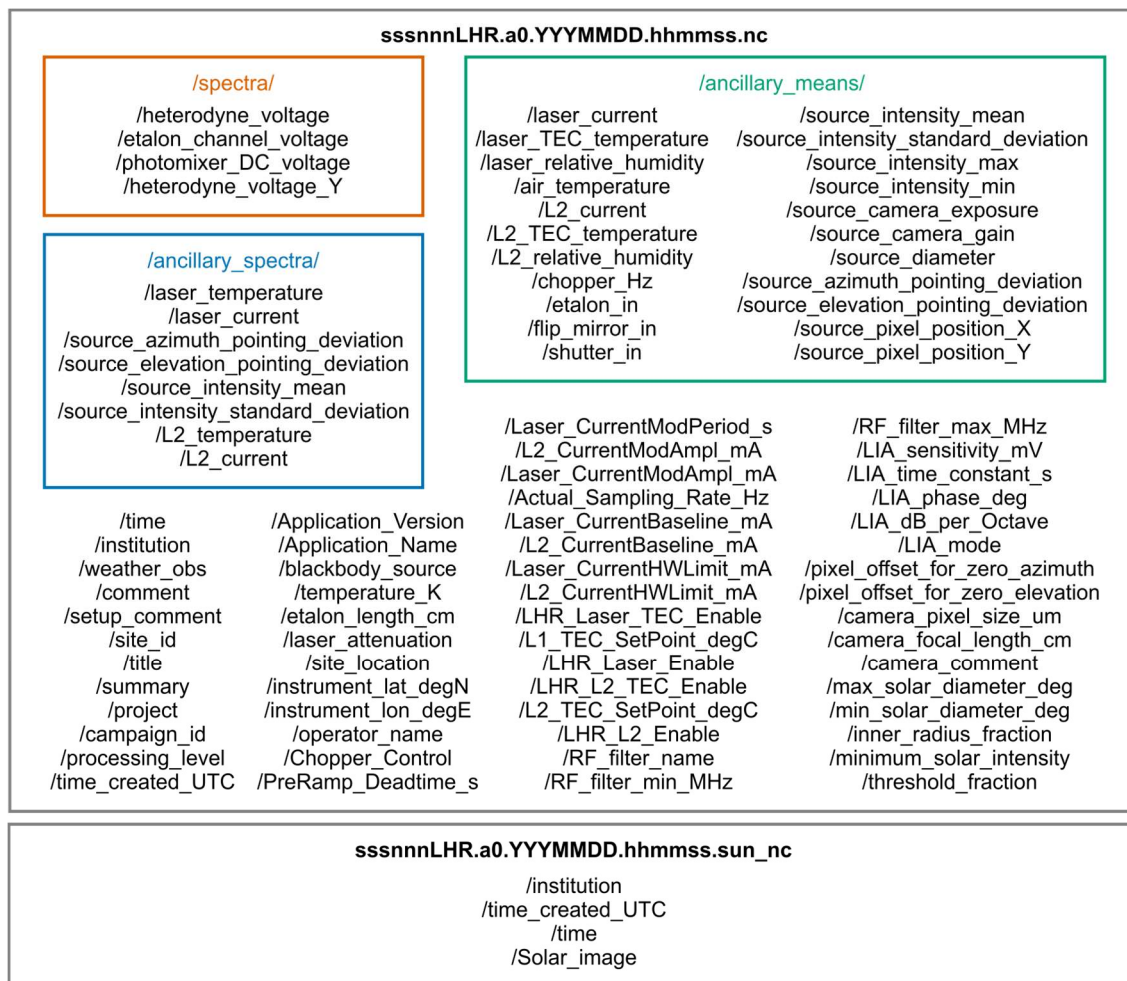


Fig. 5.2: Diagram showing the structure of the LHR (top) and HSDI (bottom) L0 NetCDF files. Coloured boxes indicate a group. Variable not inside a group are global.

5.2.2 Sun geometric elevation

The geometric solar elevation is calculated from the instrument latitude/longitude and the timestamp for every spectrum acquired. The calculation function has been validated against the NOAA solar calculator (<https://gml.noaa.gov/grad/solcalc/>).

5.2.3 Relative wavenumber calibration

The spectral relative wavenumber calibration of the laser is carried out using the reference transmission signal from a Germanium etalon (~25 mm long). From the peak and trough in the transmission signal a polynomial function mapping time into relative wavenumbers is derived with a precision of $\pm 2.10^{-5} \text{ cm}^{-1}$.

For the new processor version, the peak determination algorithm has been improved such that the measured peak position is not limited by the sampling resolution. This correction is possible because the gap between etalon peaks should change smoothly over a laser scan. In the old algorithm, a polynomial is fitted to the change in wavenumber as a function of the index points of extremums in the etalon traces. To overcome the quantization error in the index point, the new algorithm adds a small random real number to each index, effectively dithering the index numbers, and accepts the

change if the polynomial fit has improved. This is repeated until new non-integer indices are obtained that are satisfactorily fit by a polynomial. The final polynomial fit yields the improved relative frequency calibration. Whilst for FRM4GHG1 one etalon calibration trace was taken every 10 spectra recorded, the instrument has been modified to ensure 1 etalon trace is taken for every spectrum acquired.

5.2.4 Background corrections

The instrument records both the in-phase and quadrature components of the heterodyne signal during a heterodyne measurement, $I(i)$ and $Q(i)$ respectively, where i is the datapoint index. Several background signals are present in $I(i)$ and $Q(i)$ that are removed:

- A LO noise background is present due to the unwanted modulation of the laser intensity at the chopper frequency which passes through the heterodyne detection chain. This background depends on LO power, and thus varies with index i . The in-phase and quadrature backgrounds, $I_{LO}(i)$ and $Q_{LO}(i)$, are directly measured every 10 measurements by blocking the scene radiation such that only the LO illuminates the detector. Parameterized traces of $I_{LO}(i)$ and $Q_{LO}(i)$ are subtracted from the $I(i)$, and $Q(i)$ heterodyne signals during the background correction.
- A background is observed even when no LO is present, which has now been understood to originate from the modulation of the detector responsivity by variations in the scene radiant flux. The scene is constantly varying due to the amplitude modulation of the scene radiation, producing a scene background in $I(i)$ and $Q(i)$ that relates to the shot noise produced by the scene, in phase with the amplitude modulation of the scene. It scales with the total scene power contributing to a photocurrent, is fully independent of the LO power, and is therefore constant across a spectral scan. The quadrature scene backgrounds, I_s and Q_s , are directly measured when only the scene radiation is present, which occurs within the subthreshold part of every laser scan, and are also subtracted from the $I(i)$, and $Q(i)$ heterodyne signals.
- The I and Q components of the heterodyne spectrum are rotated after the above subtractions to ensure that all the transmission information is transferred into the in-phase component, yielding $I'(i)$, and $Q'(i)$. Ideally, the quadrature component $Q'(i)$ should then be zero. However, if a background contribution remains, $Q'(i)$ will exhibit a slowly varying behavior with i . The phase rotation is optimized to ensure that the de-trended quadrature component is minimised. The trend parameters are used as a residual background quality flag, to identify when the remaining background is unsatisfactory.
- Lastly, as part of FRM4GHG1 we became aware of an artifactual heterodyne background signal, evident from a non-zero signal within saturated absorption lines (Sha et al., 2024). The artifactual signal has now been understood to physically originate from the detector responsivity modulation described above, induced by the amplitude modulated scene radiant flux. As a result, the LO shot noise is partially modulated and contributes to the modulated RF power on the detection chain. The background contribution has been studied in controlled conditions and was found to scale linearly with the LO power and linearly with the total scene power contributing a photocurrent, giving a parameterized form of Eq. 5.1. It was also found to have the same phase as the genuine heterodyne signal, meaning it only needs to be subtracted from the $I'(i)$ trace.

$$I'_{\text{art.}}(i) = k \cdot V_{\text{DC}}(i) \cdot \sqrt{I_s^2 + Q_s^2} \quad (5.1)$$

The value of k depends in the detector hardware and the optical alignment of the light fields

on the detector. To establish a suitable value of k , we consider how the absorption lines depend on the solar elevation angle, θ . Assuming the atmospheric composition is constant throughout a day and ignoring atmospheric refraction, the transmittance at the centre of an absorption line is given by Eq. 5.2, where α is optical density of the absorption line and L is the atmosphere height. We find that there is a unique value of k for which the L1 processor produces an atmospheric transmittance throughout a day that follows the functional form of Eq. 5.2, as shown in Fig. 5.3. Because the CH₄ absorption line has greater optical depth than the CO₂ line, the CH₄ has lower transmittance and is thus more sensitive to changes in k . The suitable values of k were found to be $k = 10$ for the detector used in 2018 and $k = 1.5$ for the detector used in 2019

$$T = \exp\left(-\frac{\alpha \cdot L}{\sin \theta}\right) \quad (5.2)$$

A summary of the three types of backgrounds and how they typically manifest in a heterodyne measurement are shown in Fig 5.4.

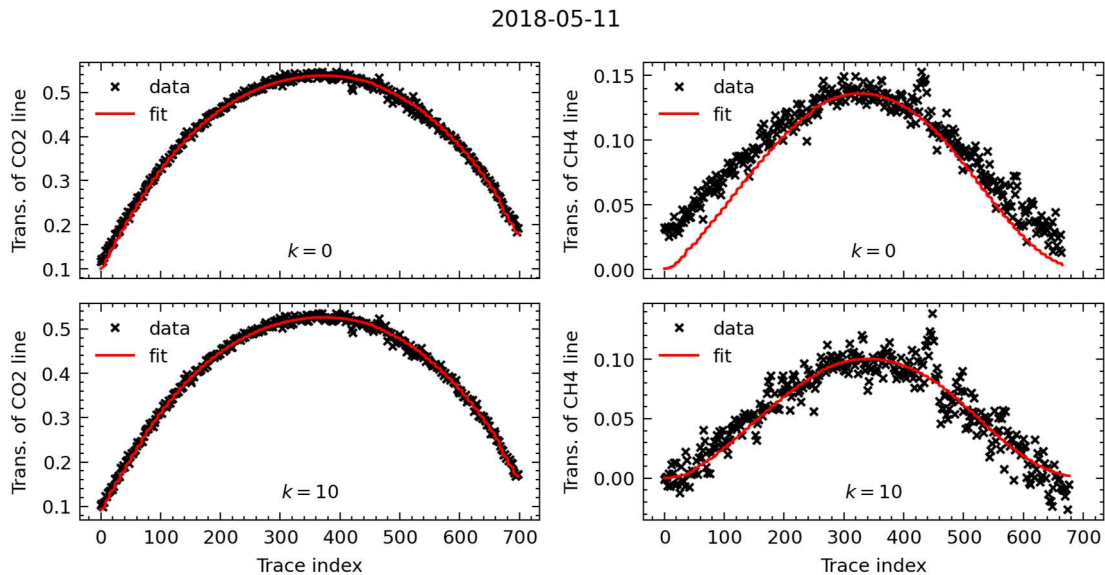


Fig. 5.3 Transmittance about the CO₂ and CH₄ lines throughout a day. Top panels show data when the L1 processor has $k = 0$, which results in a poor fit of Eq. 5.2. Bottom panels show data when the L1 processor has $k = 10$, which results in a good fit of Eq. 5.2.

SOD100LHR.a0.20190712.054609.nc, trace 475

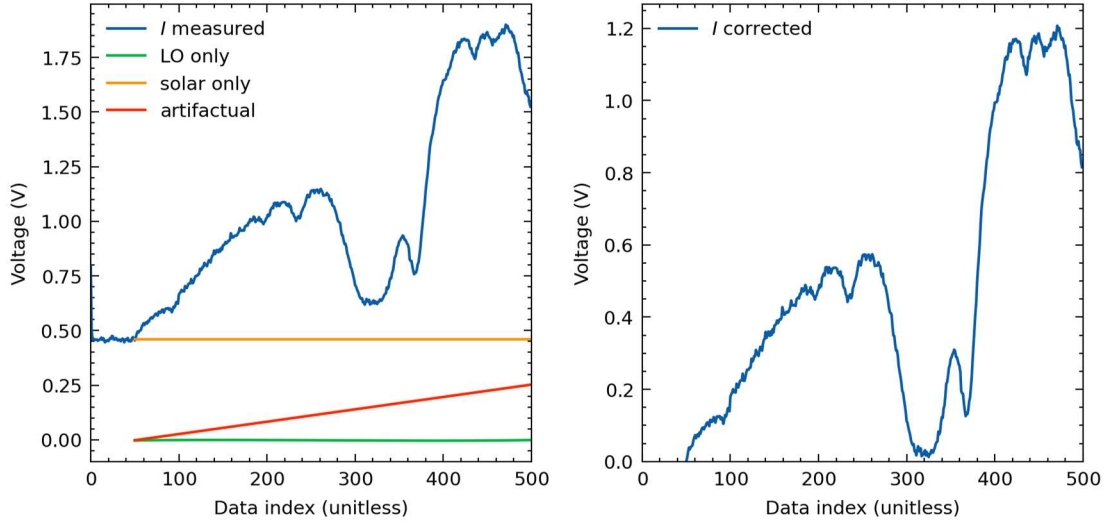


Fig. 5.4: Left plot shows a typical trace of I obtained by laser2 (ie, the CH₄ microwindow), with the three difference background contributions shown. The right plot shows the corrected I trace upon subtraction of the background signals.

5.2.5 Non-linearity correction

The non-linearity correction has not yet been implemented as it was not identified as the most significant bias source. However, this must be characterized further and we anticipate the need for a correction by leaving a place holder in the processing chain.

5.2.6 Absolute wavenumber calibration

This uses the molecular absorption features peak positions (taken from HITRAN2016 and corrected for a 1 atm pressure shift) as an absolute frequency reference. This L1 processing step has been significantly improved.

Firstly, the baseline function given in Eq. 5.4 is constructed to fit the experimental spectra as a function of relative wavenumber, σ_{rel} . It includes a polynomial term $B(\sigma_{\text{rel}}; p_0, \dots, p_n)$, a saturation curve $G(\sigma_{\text{rel}}; V_{\text{thrs}})$ and a sum of Voigt profiles $V(\sigma_{\text{rel}}; c, A, w, \gamma)$. The polynomial term describes the increase in heterodyne signal with LO power, and typically has degree one or two. The saturation curve parameterizes the gain compression of the detector, and was found to be well described by Eq. 5.4 which has the functional form of the transfer function of a second-order Butterworth low-pass filter. Gain compression onsets when the detector DC signal is above a threshold V_{thrs} ; the function $V_{\text{DC}}(\sigma_{\text{rel}})$ is a linear fit to the detector DC signal measured during the heterodyne acquisition and is fitted prior application in Eq 5.4. The Voigt profiles describe the strongest atmospheric absorption lines and are characterised by centroid c , peak optical depth A , Gaussian width w and Lorentzian width γ . All parameters after the semi-colons in the function arguments are fitted. An example a fit of Eq. 5.3 to a heterodyne trace is shown in Fig 5.5.

$$f(\sigma_{\text{rel}}) = B(\sigma_{\text{rel}}; p_0, \dots, p_n) \cdot G(\sigma_{\text{rel}}; V_{\text{thrs}}) \cdot \exp\{-\sum_{i=1}^N V(\sigma_{\text{rel}}; c_i, A_i, w_i, \gamma_i)\} \quad (5.3)$$

$$G(\sigma_{\text{rel}}; V_{\text{thrs}}) = \left(1 + \left(\frac{V_{\text{DC}}(\sigma_{\text{rel}})}{V_{\text{thrs}}}\right)^4\right)^{-1/2} \quad (5.4)$$

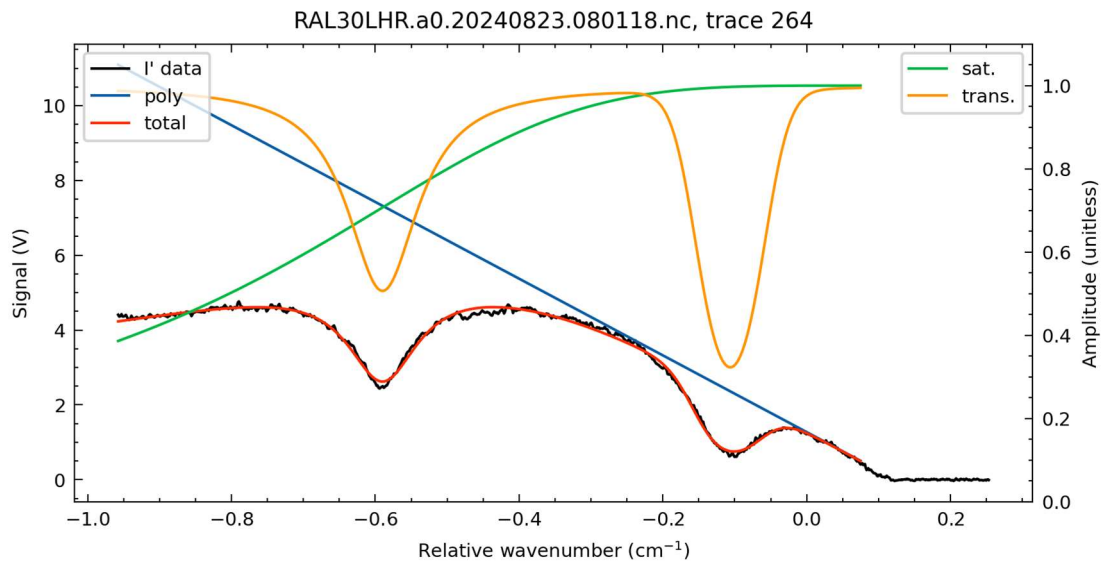


Fig. 5.5: Plot showing a typical fit of Eq. 5.3 to I' over a relevant wavenumber range. The polynomial term, saturation curve, and transmittance terms are also plotted separately.

Secondly, the difference between the fitted Voigt centroids and the known molecular absorption feature frequencies yields the absolute wavenumber correction, $\delta\sigma$, such that $\sigma_{\text{abs}} = \sigma_{\text{rel}} + \delta\sigma$. Because the spectrum exhibits more than one absorption line, the peak position difference is used to scale the wavenumber, correcting for any systematic error in the etalon free spectral range. Fig. 5.6 shows the frequency calibration error that was present prior to the scaling correction.

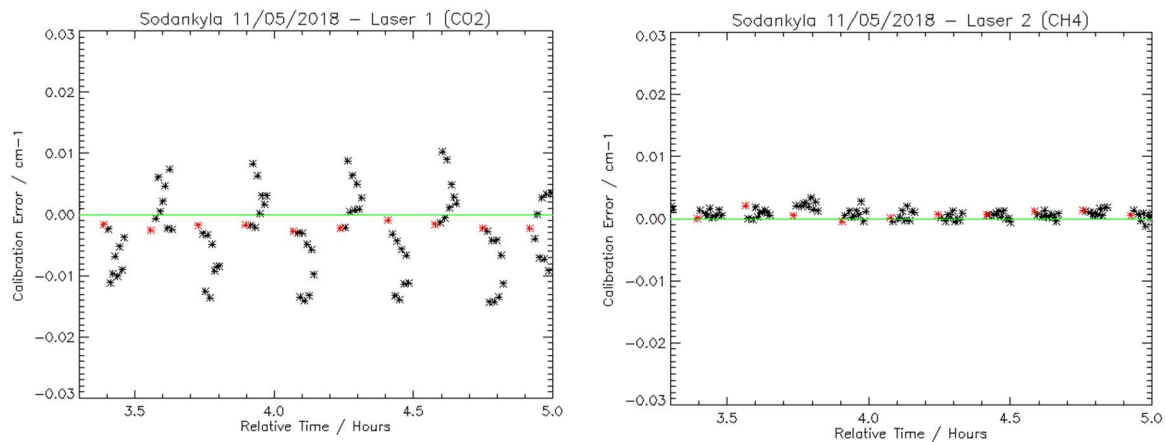


Fig. 5.6 Frequency calibration error of the old L1 processor for an excerpt of measurements taken on 11/05/2018 at FMI Sodankyla. The calibration error is corrected in the new L1 processor by scaling the frequency axis using the known molecular absorption feature frequencies. The calibration error is defined as the difference between wavenumber separation of the two fitted Voigt profiles and the wavenumber separation of the peaks taken from the HITRAN reference. Every tenth measurement, indicated by the red points, records an etalon trace resulting in a small calibration error. For the nine intermediate traces, the calibration error grows.

In preparation to pass L1 data to the L2 processor (retrieval), some extra processing is applied. The normalised wavenumber scale is constructed according to Eq. 5.5, where $\Delta\sigma$ is the spectral window width, and σ_0 the spectral window centre. This is the frequency scale used in the L2 processor. In addition, the processed heterodyne signal is divided by the gain compression function, such that the final signal of the L1 processor is given by Eq. 5.6. This division is applied because the L2 processor

assumes a polynomial baseline with atmospheric absorption features and cannot account for the non-linearity of the gain compression.

$$\tilde{\sigma} = \frac{\sigma - \sigma_0}{\Delta\sigma} \quad (5.5)$$

$$L1 \text{ signal} = \frac{1}{G(\sigma_{\text{rel}}(i); V_{\text{thrs}})} (I'(i) - I'_{\text{art.}}(i)) \quad (5.6)$$

The baseline coefficients are converted to the normalised wavenumber scale and are provided as an a priori estimate for the baseline parameters of the Level 2 processor. The residuals from the fitted function also provides a first estimate for the spectrum SNR.

5.2.7 Quality flagging

The last step of the level 1 processing generates quality flags. The approach is based on a series of threshold values contained in the processor configuration file, which will activate specific flags.

These include:

- Solar tracker quality flags: The images recorded by the HSDI can be used to characterize the quality of the Sun tracking. For each LHR spectrum, statistics on the azimuth angle error, elevation angle error, solar intensity and its variation are recorded. Either technical error or cloud contamination is identified.
- Wavenumber calibration flags: The spectral calibration sub-routines return flags indicating whether the spectral calibration error meets the threshold criteria. This includes poor quality data and cases where the fitting routines fail to converge.
- Spectral data quality flags: Flags are raised if the measured noise levels exceed a threshold, and if the fitted baseline coefficients lie outside a threshold region.

Data is only passed through to the L1B file if all quality control flags are not raised.

5.2.8 L1 data file

The L1 NetCDF has single level grouping shown in Fig. 5.7. It has a typical file size up to 10 MB.



Fig. 5.7: Diagram showing the structure of the pre-processed L1 NetCDF file. Coloured boxes indicate a group. There are no global variables.

5.3 L1 data improvement summary

A comparison of the old and new L1P data output on a typical laser1 trace is shown in Fig 5.8. The new L1 signal intensity varies by a larger amount across the trace because the detector gain compression has been divided out. The new L1 processor fits the baseline much better, accounting for the finite absorption in the wings of the absorption line. In contrast, the old processor assumed zero absorption in the region between the two absorption lines, which is invalid.

A further comparison between the old and new level 1 processor across a whole day is shown in Fig 5.9. Here, the L1 transmittance (ie, L1 signal divided by L1 baseline) in three different spectral regions of interest (ROIs) are tracked throughout the day for both old and new processors.

ROI 1 contains the edge of an absorption line, such that the mean transmittance within this ROI is sensitive to frequency shifts. The mean transmittance within ROI 1 exhibits reduced scatter for the new L1, confirming that the frequency calibration has been improved.

ROI 2 is centred on a CH₄ line that occasionally saturates, and thus is sensitivity to biases remaining after background subtraction. The mean transmittance within ROI 2 decreases to zero at the start and end of day for the new L1 data, as expected for a saturated line. Furthermore, the functional dependence of the new L1 data in ROI 2 follows the expected functional dependence upon elevation, as discussed in in section 5.2.4 and Fig. 5.3.

ROI3 is located far from the absorption lines, and thus is sensitive to weakly absorbing features. The old L1 data shows a stable transmittance of 1, suggesting the absorption in this ROI is always

negligible. However, a few percent absorption is expected within ROI 3 from the wings of the saturated line and from a weak CH₄ absorption line at 1233.96 cm⁻¹ that is 0.5% the strength of the main CH₄ line. The new L1 data for ROI 3 shows a small absorption as low elevation angles, demonstrating the expected behaviour.

SOD30LHR.a0.20180511.025408.nc, trace 500

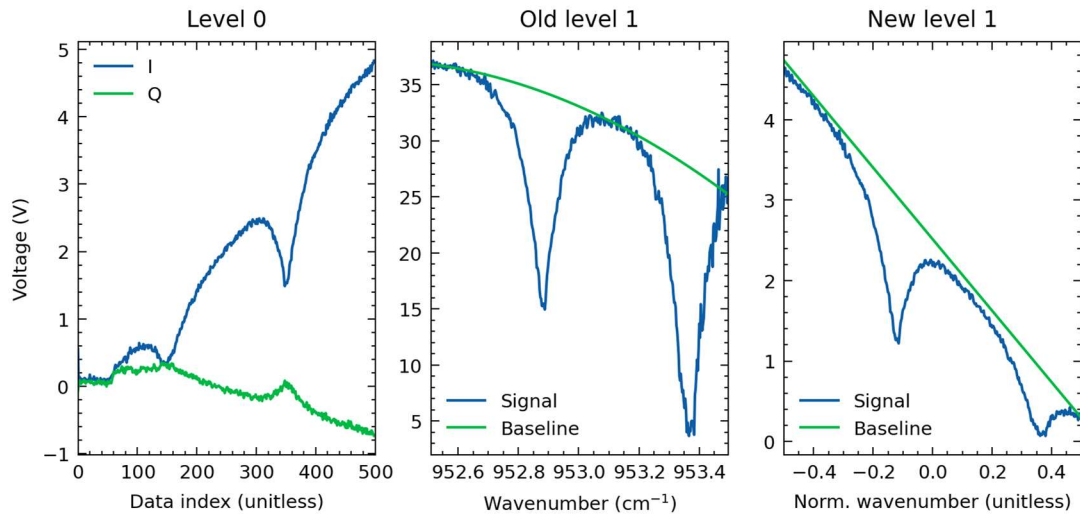


Fig. 5.8: Left panel shows the unprocessed I and Q traces of a typical laser1 trace from a L0 dataset. The middle panel show the L1 signal and baseline from the old level 1 processor, whereas the right panel shows the same for the new level 1 processor.

2018-05-11, laser 2 (CH₄ microwindow)

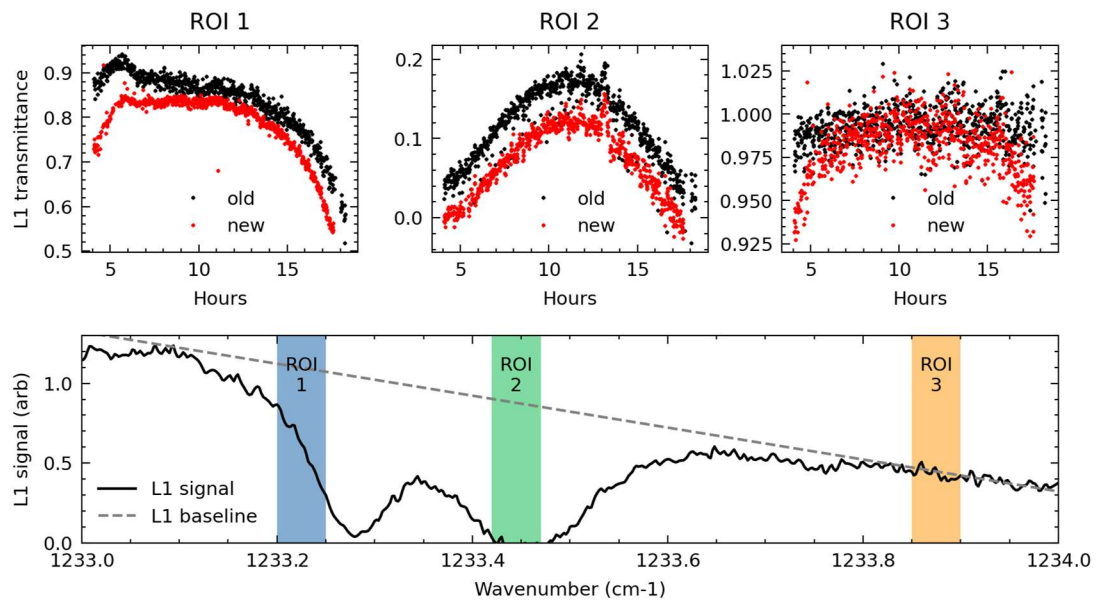


Fig. 5.9: L1 transmittance across one day of measurements with laser 2. Bottom panel show the new L1 signal and baseline of an example trace, and highlight the three spectral regions of interest (ROI). The left, middle and right panels on the top show how the transmittance changes across the day, for both old and new L1 processors.

6 LEVEL 2 PROCESSOR

6.1 L1C pre-processor

The L1C pre-processor transforms the L1B files (NetCDF) into L1C files (ascii) ready for ingestion into the level 2 processor.

The L1B file contains separate timeseries of laser1 (CO₂ microwindow) and laser2 (CH₄ microwindow) scans. The L1C pre-processor then:

- Constructs a single time-ordered list of the L1B data by merging the individual series of laser1 and laser2 data.
- Identifies in the list when a sequence of laser1 plus laser2 scans is completed (i.e. when the instrument switches back to laser1) - these define the full scans number of a particular laser – and store the positional index of these.

The next step is for the pre-processor to remove some of the baseline fluctuations. The set of samples consists of N experimental spectra from a given laser $\{S_1(\sigma), \dots, S_N(\sigma)\}$. Over the duration of measurement for the full set, the atmospheric transmittance $\tau(\sigma)$ is assumed constant. Therefore, each spectrum i can be written as $S_i(\sigma) = \tau(\sigma) \cdot B_i(\sigma)$ where $B_i(\sigma)$ is a quadratic baseline introduced by the measurement system.

The averaged spectrum $\bar{S}(\sigma)$ from the set is calculated. Given the assumption made on the transmittance, we have $\bar{S}(\sigma) = \frac{\tau(\sigma)}{N} \sum_{i=1}^N B_i(\sigma)$.

Then, for each spectrum i , the parameters of $B_i(\sigma)$ are re-adjusted so that $\bar{S}(\sigma) - S_i(\sigma)$ is minimized. The new set $\{S'_1(\sigma), \dots, S'_N(\sigma)\}$ is then averaged as the L1C spectrum.

Lastly the ascii L1C data file is written. In addition to the collection of spectra and their associated noise vector, the L1C contains the corresponding geolocation, time/date stamps, and solar elevation angles for each trace.

6.2 Multispectral Orbital Retrieval using Sequential Estimation

The level 2 processor is the “retrieval” algorithm. We have made a complete change of the retrieval software and moved from a specifically written optimal estimation algorithm (Weidmann et al., 2007) to a more generic one inherited from the MIPAS mission and initially developed in tandem with RFM as the Multispectral Orbital Retrieval using Sequential Estimation (MORSE - <https://eodg.atm.ox.ac.uk/MORSE/>) and programmed in the FORTRAN language by Anu Dudhia (University of Oxford).

MORSE has been modified in order to implement retrievals in the ground-based zenith solar occultation viewing mode, and is referenced as MORSE-GB hereafter.

MORSE uses a ‘sequential estimation’ method which is a particular implementation of the standard optimal estimation iterative solution.

$$x_{i+1} = x_i + [\gamma S_a^{-1} + K_i^T S_y^{-1} K_i]^{-1} [K_i^T S_y^{-1} (y - f(x_i)) - S_a^{-1} (x_i - a)] \quad (5.5)$$

where x is the state vector; subscripts $i, i + 1$ are iteration numbers; y is the measurement vector; a is the a priori estimate of x ; f is the forward model (RFM is used); S refers to the covariance matrices of the various vectors x, a, y ; and γ is the Levenberg-Marquardt parameter.

The ‘sequential’ aspect, as implemented in MORSE, is not to solve the equation using all available measurements simultaneously but to divide the measurements into correlated subsets, and solve

for each in turn, updating the a priori estimate after each.

A χ^2 statistic is defined before the sequential update begins $i = 0$ and again after each iteration. x_0, S_{x0} refer to the a priori estimate for the current update rather than the (climatological) initial values used at the start of the retrieval. This is the ‘cost-function’ that the retrieval attempts to minimise through each iteration.

$$\chi_i^2 = (x_i - x_0)^T S_{x0}^{-1} (x_i - x_0) + (y - f_i)^T S_y^{-1} (y - f_i) \quad (5.6)$$

One of the main differences compared to the MORSE implementation for limb solar occultation (Weidmann et al., 2020), relates to the absence of exo-atmospheric reference allowing the calculation of atmospheric transmittance and the need to handle a slowly varying baseline modulating the transmittance spectra.

To account for this effect, the parameter to be fitted is the transmittance adjusted for baseline scaling. It takes the form given in Eq. 5.7. This form was retained as the baseline parameters a priori are all set to zero.

$$\tau'(\tilde{\sigma}) = \tau(\tilde{\sigma}) \cdot (1 + a_0) \cdot (1 + a_1 \cdot \tilde{\sigma} + a_2 \cdot \tilde{\sigma}^2) \quad (5.7)$$

6.3 Inputs

The inputs to MORSE-GB include:

- The L1C file generated from the L1B by the pre-processor (section 6.1).
- The a-priori climatology, currently taken as the NCEP files generated for Sodankylä, as well as a WACCAM CO₂ profile,
- The HITRAN database (the 2016 database was used).
- The instrument lineshape function.

For the reprocessing, the local surface pressure, temperature, and relative humidity measurements were not yet used to constrain the climatology a priori-data. The field of view of the instrument was taken as a pencil beam (no field of view function).

6.4 Preliminary example run

At the time of writing, the full L2 retrieval has not been finalized nor consolidated. We are giving here an excerpt of the very first runs, which do require more critical analysis and work.

The illustration from Fig. 6.1. is taken from the 11/05/2018 granule. The Shannon information from the retrieval was typically between 12 and 15 for the three molecules retrieved: CO₂, CH₄ and H₂O.

For these L2P runs the scaling parameters given in Eq. 5.7 were taken to be zero with a standard deviation of 0.1.

Based on the early runs, we suspect remaining negative biases for both XCO₂ and XCH₄, which seems on the lower than the global averages. The existence of biases is also apparent in the water column retrieval, as water is retrieved from both the CO₂ and the CH₄ micro-windows, they ought to match but a time dependent bias is perceptible in the record (see Fig. 6.1).

The Fig. 6.1 also shows the equivalent day from the historical data, which were denser as no L1C pre-averaging was done. The full impact of this pre-averaging still needs to be studied.

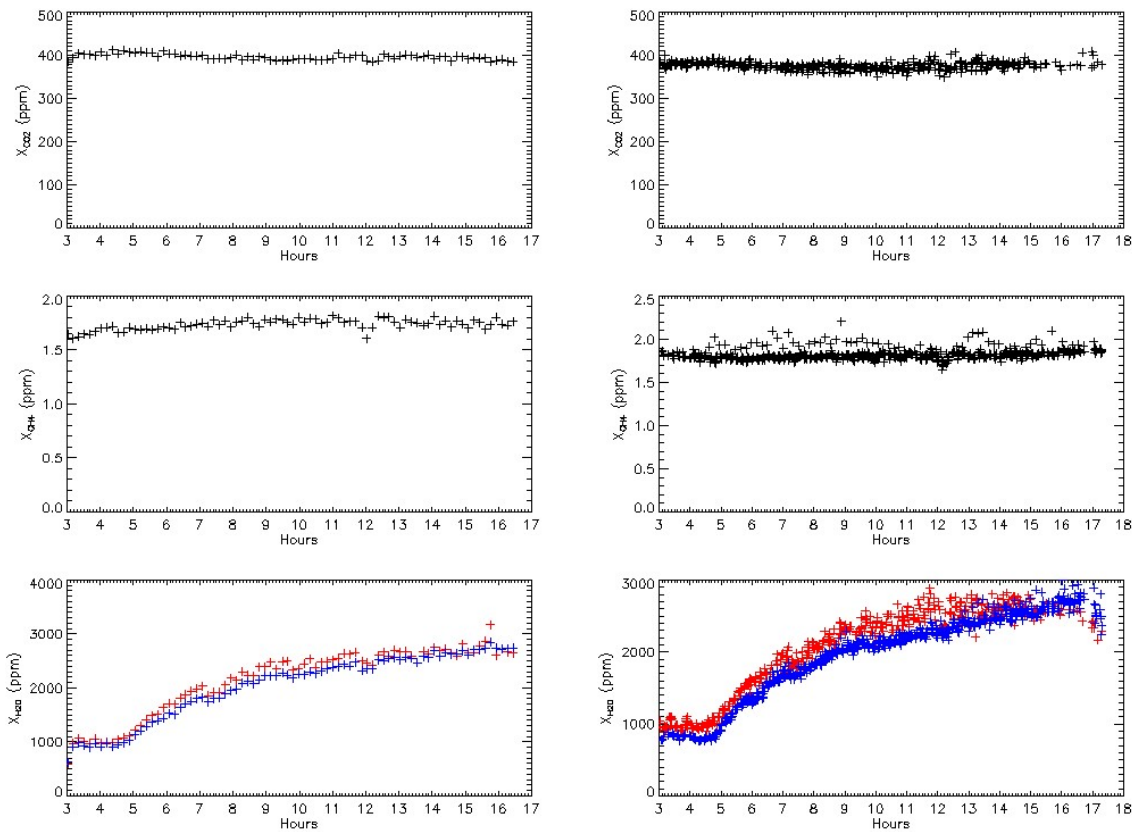


Fig. 6.1: L2 data set extracted for the 11/05/2018 data file. The plots shows the diurnal evolution the retrieved column of XCO₂, XCH₄m and XH₂O as function of time of the day in hours (UTC). The lower panel on XH₂O shows the water vapour column retrieved from the CO₂ channel (red) and the CH₄ channel (blue), suggesting some residual biases evolving during the day. Left, is the output from the new L2P, right is the old dataset.

7 CONCLUSION AND FORWARD LOOK

7.1 Full processor update

Through the LHR instrumental overhaul and laboratory work, the knowledge in biases physical origin has progressed significantly. The level 1 data processor has been updated accordingly to account for improved background corrections, improved I and Q rotation, and baseline estimation. In addition, the previous wavenumber calibration was found to be inaccurate and has been improved.

As a result, a novel L1 dataset from the historical Sodankyla data was generated. Qualitative assessment based on expected diurnal trends suggests the data quality and physical representativity has been improved.

The level 2 data processor has been completely overhaul into a new sequential optimal estimation FORTRAN code. The ancillary data input files and pre-processor have been developed to allow for ingestion of the FRM4GHG dataset from the historical data and the forthcoming ones.

The new L2P has only recently produced its first output, therefore it runs. The corresponding Level 2 dataset has been produced, but should be taken with caveat as the Level 2 configuration and parametrization is still incomplete at the time of writing.

7.2 Future instrumental and operational changes

Beyond the reprocessing of the historical dataset from 2018-2019, analysis of the L0, L1 and L2 data has identified several biases that can be eliminated or reduced through hardware changes or changes to the instrument operation. These changes made to the instrument, relevant to measurements acquired in 2024 and beyond, are outlined below.

- The retrieved total columns were correlated with air temperature within the instrument, suggesting that the measurements are biased by instrument temperature. To reduce the temperature bias on future measurements, the thermal stability of the instruments has been improved by installing insulation on all walls and installing cold plates on the instrument floor that makes use of the water-cooling system already used to remove excess heat from the thermoelectric coolers.
- The background signals discussed in Section 5.2.3 were identified to originate from modulation of the detector's responsivity, induced by the modulated scene radiance. To reduce the background signals in future measurements, a narrow optical bandpass filter centred on the CO₂ line of interest is now placed in the scene's beam path during a CO₂ measurement. This was found to reduce the background signal in the CO₂ measurement.
- A laboratory characterization identified that the lock in amplifier's gain depended on the modulation frequency. To eliminate changes in gain due to drifts in the modulation frequency, the frequency of the chopper is now controlled via a feedback loop and kept stable to a standard deviation of 4 Hz.
- Analysis of the frequency stability of the instrument, shown in Fig. 5.6, show the laser frequencies drift significantly between each scan. To reduce this effect on retrievals, heterodyne measurements are now taken with the etalon always present in the reference detector's beam path, rather than every 10th laser scan. To avoid the presence of the etalon degrading the signal-to-noise of the heterodyne signal, the instrument hardware was modified to include an isolator made of a polarizer and quarter-wave plate for each laser.
- The Thorlabs infrared CCD previously used to image the solar disk has been replaced by a Ximea hyperspectral camera that has sixteen spectral channels from 590 nm to 860 nm. This captures spectroscopy data complementary to the LHR. Because the solar image now contains spectroscopy data, the L0 data files have been restructured accordingly. There is now one NetCDF file for LHR data, another for the hyperspectral images, and another for the solar tracker. The L0 structure of the LHR data, shown in the top panel of Fig. 5.7, has been modified to remove solar tracker information and restructured to reduce the number of global variables.

Furthermore, we plan to investigate whether k , the parameter that dictates the strength of the artifactual background, can be measured empirically in real time. This would eliminate biases introduced from the assumption used to derive Eq. 2.2.

8 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

9 REFERENCE DOCUMENTS

- Dudhia, A., 2017. The Reference Forward Model (RFM). *J. Quant. Spectrosc. Radiat. Transf.* 186, 243–253.
- Sha, M.K., De Mazière, M., Notholt, J., Blumenstock, T., Bogaert, P., Cardoen, P., Chen, H., Desmet, F., García, O., Griffith, D.W.T., Hase, F., Heikkinen, P., Herkommer, B., Hermans, C., Jones, N., Kivi, R., Kumps, N., Langerock, B., Macleod, N.A., Makkor, J., Markert, W., Petri, C., Tu, Q., Vigouroux, C., Weidmann, D., Zhou, M., 2024. Fiducial Reference Measurement for Greenhouse Gases (FRM4GHG). *Remote Sens.* 16, 3525.
- Weidmann, D., Antonini, K., Martinez Pino, D., Brodersen, B.K., Patel, G., Hegglin, M.I., Sioris, C., Bell, W., Miyazaki, K., Alminde, L.K., Gabriele, A., Pastena, M., Hoffmann, A., 2020. Cubesats for monitoring atmospheric processes (CubeMAP): A constellation mission to study the middle atmosphere. In: *Proceedings of SPIE - The International Society for Optical Engineering*.
- Weidmann, D., Reburn, W.J., Smith, K.M., 2007. Retrieval of atmospheric ozone profiles from an infrared quantum cascade laser heterodyne radiometer: results and analysis. *Appl. Opt.* 46, 7162.