

REPORT ON THE NOMAD SO INSTRUMENT LINE SHAPE (ILS) CALIBRATION

Giuliano Liuzzi and Geronimo L. Villanueva (NASA Goddard Space Flight Center)

May 2021

Table of Contents

| | |
|---|----|
| 1. Preface | 1 |
| 2. Instrumental line shape basics | 2 |
| 3. ILS calibration and modeling | 3 |
| 4. Functional form of the ILS and its origin..... | 5 |
| 5. Results: examples for a single order and full solution | 6 |
| 6. Comparison with other ILSs: BIRA, IAA/CSIC (as of Mar/2021)..... | 8 |
| 7. Impact of the ILS on retrievals: are A_2 and σ uniform? | 10 |
| 8. Retrieval of full CO_2 profile with the new ILS solution | 12 |
| 9. The origin of the offset | 13 |
| 10. Final considerations: ILS recipe and offset recommendations | 14 |

1. Preface

The following documents describes updated results and efforts to produce a new calibration recipe for the Instrumental Line Shape (ILS) of NOMAD SO and a corresponding updated model for the spectral offset, which has been observed since the first calibration described in [Liuzzi et al. \(2019\)](#). The results obtained in this report are obtained by including a larger amount of data and orders than previous analyses, and by fully integrating the ILS derivation with the spectral offset, which ultimately is attributed to the AOTF.

The new ILS results show that the recipe derived in May/2020 at GSFC is very similar to the new one, as shown in Figure 1. The new resolving power is 15% lower than the one derived initially, while the difference in the relative amplitude of the ILS components is only 0.05. All these differences can have a variable impact on different diffraction orders and different altitudes. Individually, the relative difference between two retrievals made with the two sets of ILS parameters should not be larger than the relative differences between the ILS parameters themselves (20%). Previous retrievals of water vapor, D/H and CO are done by mixing the information coming from several orders at each altitude, which blends the effect of the ILS parameterization with other sources of uncertainty of similar magnitude. Other retrieved parameters that do not require high-resolution radiances, such as aerosols, do not depend on the ILS parameterization.

As for the AOTF, the offset has been investigated in previous works and widely adopted in previous retrievals as well and was able to grant satisfactory results in terms of consistency among different orders and spectral residuals. In this work, we introduce a new method to obtain an estimation of the offset, whose order of magnitude is consistent with previous works.

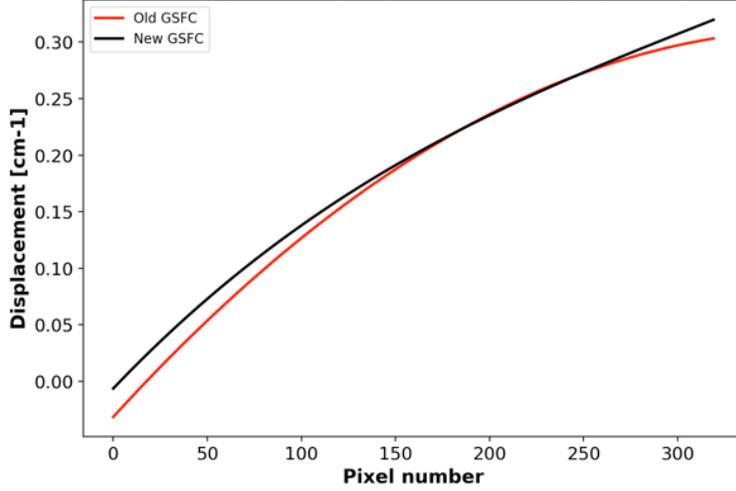


Figure 1. Displacement between the two gaussians describing the ILS according to the original GSFC recipe (May/2020) and the new one derived in this report.

2. Instrumental line shape basics

When the radiation is received by an instrument, one or more optical components of the instrument (e.g. the slit, the grating, the geometry of the optical path of the light) result in the final spectral resolution. The Instrument Line Shape (ILS) of an instrument is the function that quantitatively describes the radiation as observed in each spectral channel (from now on, “pixel”) with respect to the input radiation.

In mathematical terms, let $R(\nu)$ be the radiation entering the instrument. The observed spectrum $y(x_i)$, $i = 1 \dots N$, in each of the N pixels, will result from the convolution of $R(\nu)$ with the ILS:

$$y(x_i) = \int_{-\delta\nu}^{\delta\nu} R(\nu(= x_i) + x) \cdot ILS(x) dx \quad (1)$$

As the ILS and the entering flux are described at a finite resolution, this process can be reduced to its discrete form:

$$y(x_i) = \sum_{j=j_0-n}^{j=j_0+n} R(\nu_j) \cdot ILS(j) \quad (2)$$

where the ILS kernel is defined on an appropriate number n of points and is normalized to 1 to preserve the incoming flux:

$$\sum_{j=j_0-n}^{j=j_0+n} ILS(j) = 1 \quad (3)$$

The functional form of the ILS depends on the instrument principles, configuration, and optical scheme. In the case of NOMAD, it is expected that the presence of a grating as central element of the spectrometer results in an ILS that follows the functional form of a sinc, which would yield significant contribution to the flux in a pixel with radiation from several wavenumbers away.

However, this assumption can be verified by directly deriving the ILS functional form from the data, which can be done using different methods. In this report, we describe the derivation of the ILS using a self-convolution method, which exploits the full current knowledge of the instrument response and the information content of the observed spectra.

3. ILS calibration and modeling

The self-convolution method uses the information content of a spectrum in a completely agnostic way, i.e., without making any a-priori assumption on the functional form of the ILS. Importantly, in the case of NOMAD, the AOTF plays an important role in introducing flux from nearby diffraction orders into the observed spectrum, often resulting into superposition between atmospheric lines that fall in the same spectral pixel from distant wavelengths. The self-convolution method compares the observed spectrum with a fiducial model that considers the AOTF properties, therefore it can naturally take into account this phenomenon at least to a certain degree of accuracy defined by our knowledge of the AOTF response itself.

By following the same formalism as the previous equations, the ILS can be described by the convolution between the low-resolution (=observed) spectrum and the corresponding high-resolution (=modeled) one:

$$ILS(j) \cong \int_{-\infty}^{+\infty} y(x_i + \delta)R(v)d\delta \cong \sum_{i=-M}^{+M} \tilde{y}(x_{i+j}) \cdot R(x_i) \quad (4)$$

where M represents the number of pixels of the observed spectrum around the i -th pixel of the observed spectrum, and \tilde{y} indicates the spectrum previously interpolated to the same sampling as the model radiance R to perform the sum.

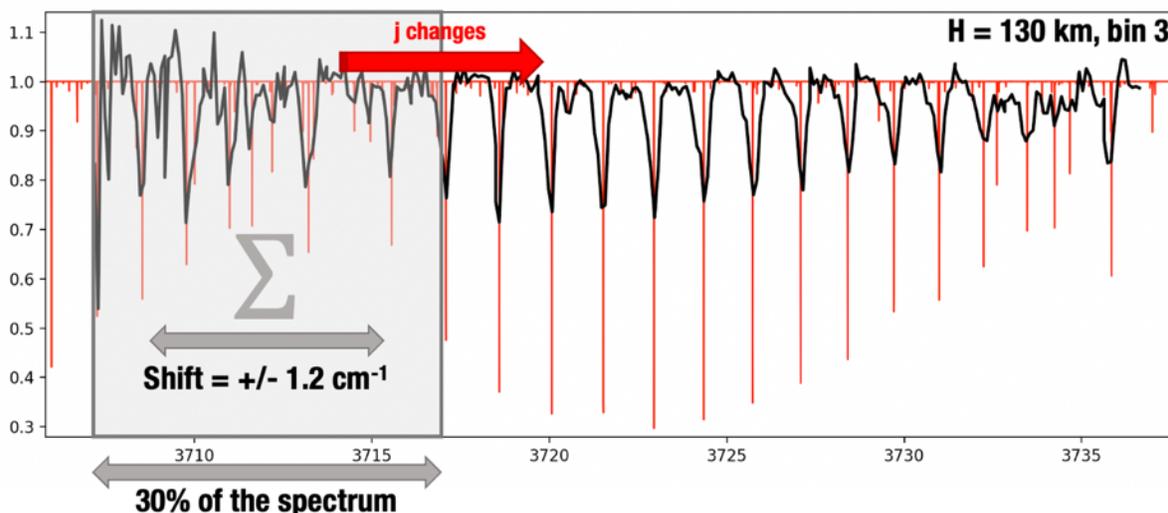


Figure 2 . In this example, the black is the observed spectrum ($\times 10$ for visibility) in order 165, while the red is the model corresponding to the best fit (which is done before the convolution, see later). Within a window corresponding to a certain portion of the spectrum (30% in this case, which encompasses M pixels of the interpolated spectrum), the observed spectrum is translated, and the correlation is performed again, building the ILS function. The width of the window (M) and the maximum shift (j) can be chosen and varied on a case-by-case basis.

As the observed spectrum is translated with respect to the model, Eq. (4) computes the j -th value of the ILS function across wavelength. A scheme for this process is shown in Figure 2 on an actual NOMAD spectrum from order 165.

When using this method, it is crucial to choose appropriately both the width of the window on which the convolution is performed, and the maximum shift to be applied to the observed spectrum. The first one defines how much signal is entered into the calculation of the ILS and should be chosen as a compromise between the need of having a characterization of the ILS as much as local as possible, and the availability of bright lines in any given point of the chosen interval.

As i is varied (Eq. (4)) across the whole observed spectrum, the ILS is reconstructed for each pixel of the observed spectrum. We can therefore summarize the main steps to estimate the ILS for all the pixels in the following steps:

- 1) For the chosen order, an observed spectrum with visible atmospheric lines is chosen.
- 2) Assuming a realistic model for the AOTF, a retrieval is performed to retrieve the abundances of the relevant molecular species, to correct and remove any continuum effect (i.e., the resulting spectrum is normalized to 1), and calibrate the wavelength of each pixel, resulting in y .
- 3) The abundance of the molecular species is used to compute the synthetic spectrum R (including AOTF contributions) at a resolving power much higher (x10, x20) than the observed spectrum. This represents the best estimate of the actual radiation reaching the grating of the instrument. Importantly, the spectral grid on which R is computed must be in the proper unit of constant resolving power, as expected for a grating spectrometer.
- 4) y is interpolated to the grid of R , and from both the continuum is subtracted.
- 5) Defining M and the maximum shift (value of j), the convolution is performed according to Eq. (4) for each original i . Typical values for these parameters can be those in Figure 2.
- 6) ILS for each pixel is normalized to 1, continuum-subtracted and stored.

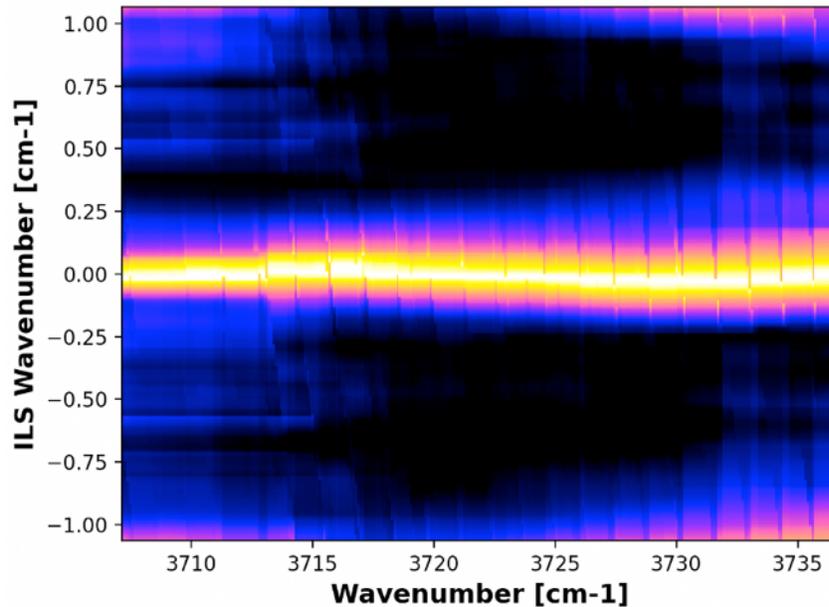


Figure 3. Estimated ILS for each spectral pixel (x -axis). The intensity of the color scale is from 0 (black) to 1 (white).

An example of estimated ILS is shown in Figure 3. The matrix shown here has the spectral pixels of the observed spectrum and their wavenumber on the x-axis and the ILS wavenumber on the y-axis. By extracting the i -th column of this image, we can see the behavior of the estimated ILS for the i -th pixel of the observed spectrum.

It can be seen that the center of each column varies slightly, with oscillations that can be as large as $\pm 0.05 \text{ cm}^{-1}$: these variations are not necessarily related to possible variability of the ILS itself, rather to residual issues in the frequency calibration of the observation. The matrix also shows aliases of nearby lines (CO_2 in this case), which are present whenever the chosen maximum j to perform the correlation is close to the separation between nearby lines.

4. Functional form of the ILS and its origin

Once the ILS is derived from the self-correlation, the next step consists into elaborating an a-priori model to fit to it. While, as said, in principle the ILS of a grating is a sinc function, in many cases the line core can be reasonably modeled by a gaussian kernel which will be mostly defined by the resolving power of the instrument.

In the case of NOMAD SO, the data suggest that the most probable form for the ILS is given by the superposition of two gaussian kernels with variable separation across the order. To be more specific, the observed spectrum has been seen to result from the superposition of two identical images of different intensity, each one resulting from a single gaussian kernel. Using the previous formalism, we can then define the spectrum R_1 from the convolution with the first ILS, and the spectrum R_2 from the convolution with the second ILS, defined as:

$$\begin{cases} ILS_{1,i}(x) = A_{1,i} \exp\left[\frac{-(x - \nu_i)^2}{2\sigma_{1,i}^2}\right], i = \{0, \dots, 319\} \\ ILS_{2,i}(x) = A_{1,i} \cdot A_{2,i} \exp\left[\frac{-(x - \nu_i - b_i)^2}{2\sigma_{2,i}^2}\right], i = \{0, \dots, 319\} \end{cases} \quad (5)$$

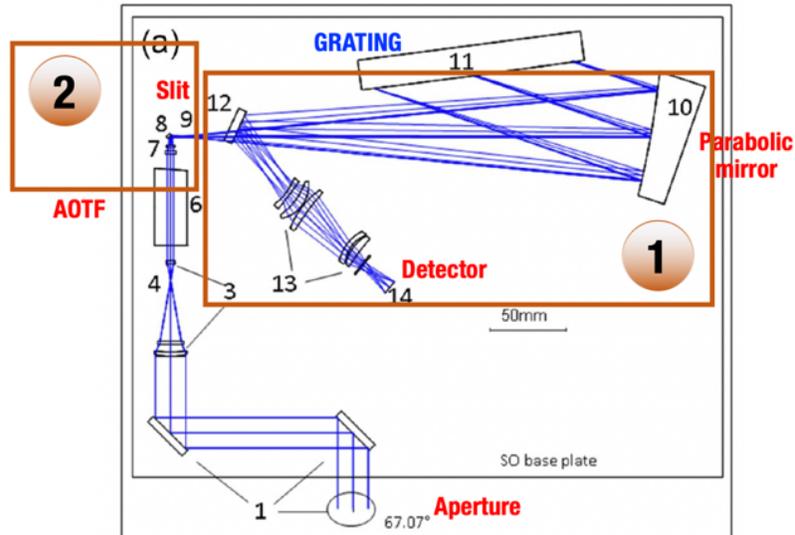


Figure 4. 2-D optical scheme of NOMAD SO. The region 1 includes all the optical element that the light encounters after being diffracted by the grating, while the region 2 includes the AOTF and the slit.

The origin of this ILS form is intimately connected to how each of the parameters in Eq. (5) relate to each other. A schematic view of this is shown in Figure 4, with the optical scheme of NOMAD SO and the possible origin for the double gaussian ILS.

If the double gaussian is generated at any point of the region 1 by an optical element (e.g., a cracked lens, a slight misalignment), both the gaussians must preserve the properties dictated by the grating (and previously the slit) characteristics, that is essentially the width. In case the double gaussian arises somewhere in the region 2, e.g., because of the AOTF or the slit, that would still result in different images of the source that have to go through the grating, which would ultimately yield the two images with the same resolution.

Based on this, we choose to simplify the parameterization given in Eq. (5), assuming that $\sigma_{1,i} = \sigma_{2,i} = \sigma_i$ for every pixel. In addition, we choose to give ample a-priori liberty to the relative shift b_i between the two gaussians, describing it as a third-order polynomial of the pixel number.

The other assumption (verified later) about this parameter is that, if the second gaussian ILS follows the grating equation, the shift must do the same, and must be somehow scalable with the wavelength, that is with the diffraction order. Therefore, when we derive the ILS functional form and the coefficients, we do that for single orders, and then we verify whether or not this hypothesis holds.

In synthesis, the form of the coefficients we choose as representative of the ILS is the following, for a total of 6 coefficients:

$$\left\{ \begin{array}{l} A_{1,i} = A_1 \quad \forall i \\ A_{2,i} = A_2 \quad \forall i \\ b_i = poly(\beta_{0\dots3}, i) \\ \sigma_i = \sigma \quad \forall i \end{array} \right. \quad (6)$$

5. Results: examples for a single order and full solution

Figure 5 shows how the derived ILS for order 165 looks like, the corresponding best fit (as an image) and some examples of extracted rows, corresponding to specific spectral pixels. The spectrum that has been used for this exercise is acquired at 130 km, where the lines are not expected to be significantly saturated, if at all.

The solidity of this solution can be tested by trying to investigate how it varies by changing the observed spectrum and moving to altitudes where spectral lines are saturated. This is a key test, as other methods can be susceptible to line saturation, which could result into a bias in the apparent line width and resolving power. Figure 6 shows how this can impact the ILS parameters derived at different degrees of saturation. Despite using saturated lines, the self-correlation is still able to yield consistent results across different regimes.

Once a solution is derived for a single order, we repeat the procedure for all the orders. For the ILS derivation, it is best practice to use orders across the whole spectral range covered by NOMAD and to focus on those orders where lines are of similar intensities and with spacing as much as regular as possible. Based on this, we have chosen to work on the following orders: 119, 121, 124, 140, 141, 145, 149, 154, 156, 161, 165, 171, 190, 191. To work on a uniform dataset, we have chosen spectra from the atmospheric full-scan 20181006_090808, which comprises ~ 15 acquisitions per order per spectral bin at altitudes from 0 to 200 km. In this report, we consider only spectra from bin 3, which has been used at GSFC in routine retrievals. Yet, because of the suspected cause of the double gaussian shape for the ILS, it is quite possible that the solution we obtain here could be not adaptable to other spectral bins, because of projection effects onto the detector.

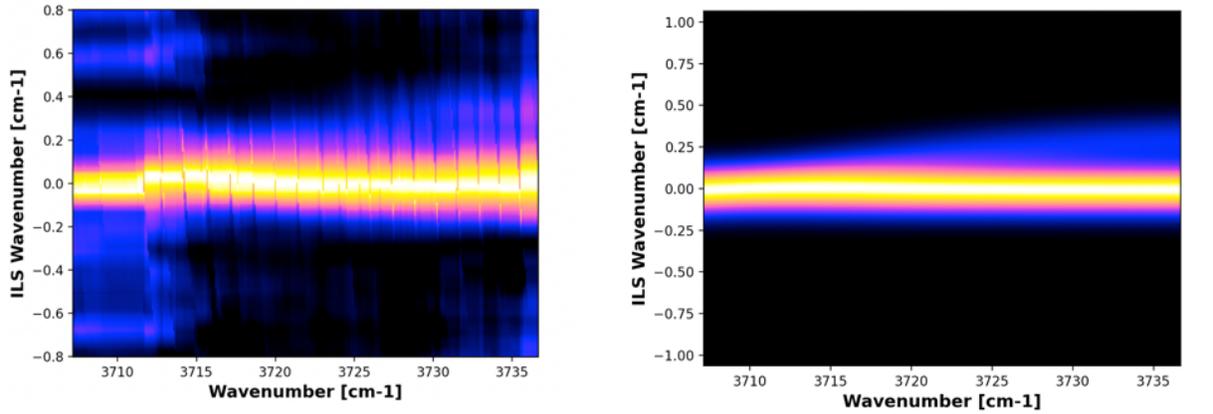


Figure 5. ILS solution for order 165. Left top: derived ILS. Right top: ILS best-fit model according to the previous equations. Bottom panels: extracted columns from the ILS image (black) and best fit (red) corresponding to pixels 20, 145 and 295 out of 320. X-axis is purely indicative, as it is in wavenumbers \times pixel number.

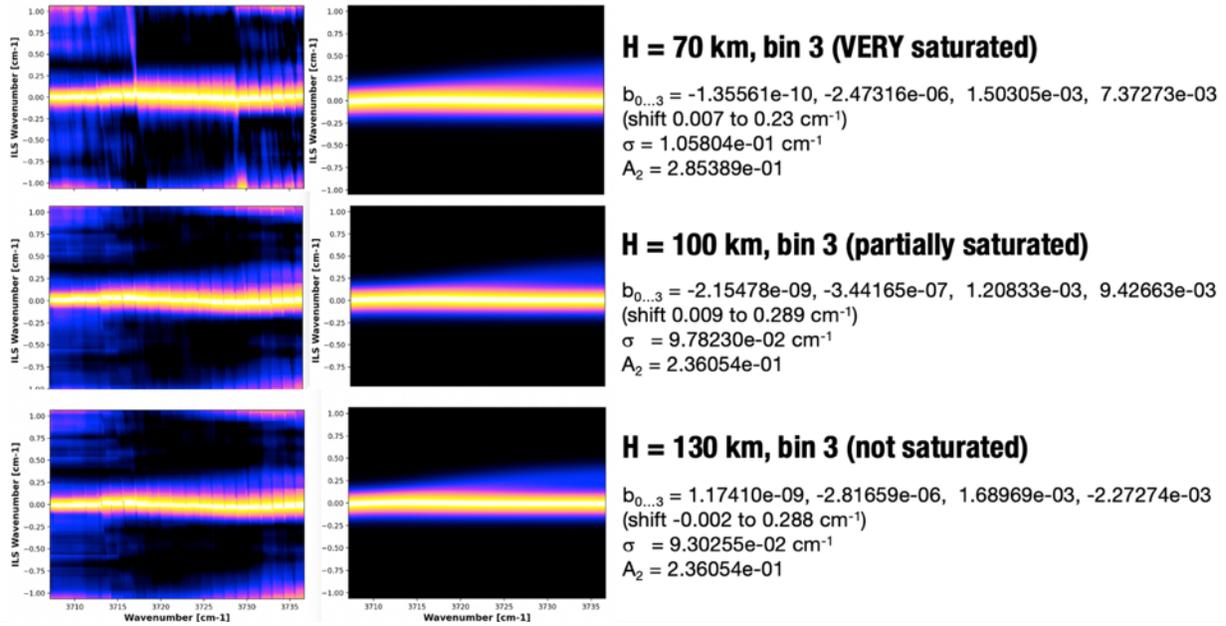


Figure 6. Left column: derived ILS for 3 different altitudes characterized by different saturation regimes of the spectral lines. Middle column: corresponding best fit model for the ILSs. Right column: retrieved values for the coefficients in Eq. (6) and corresponding boundaries for the shift and RP.

For all the orders, we have chosen one single spectrum at an altitude where the H₂O, CO₂, or CO lines are not saturated, that are respectively: 26 km for orders 119, 121 and 124; 42 km for orders 140, 141 and 145; 59 km for order 149; 80 km for orders 156 and 158; 130 km for orders 161, 165; 66 km for order 171; and 85 km for orders 190 and 191.

The polynomial fit for the displacement between for the two gaussians with respect to the pixel number is shown in Figure 7. Importantly, in this plot the displacements are normalized by the ratio between a reference wavenumber (3700 cm⁻¹) and the central wavenumber of the order. The fact that the normalized curves are consistent with each other shows that the displacement follows the grating equation, where the pixel pitch depends (approx. linearly) on the wavelength.

The final expression, common to all orders, is obtained by a final least-squares minimization of the polynomial expressions obtained by each order and is shown in Figure 7 in solid black.

As for the other parameters, i.e., the resolving power and the amplitude of the second gaussian with respect to the first, we find a constant value of respectively 17,000 and 0.30 across all the orders. In a next section, we will examine the possible variations of these two parameters.

6. Comparison with other ILSs: BIRA, IAA/CSIC (as of Mar/2021)

The solution we have retrieved for order 165 has been tested and compared to the current (Mar/2021 latest update) solutions from other approaches, which are mostly based on fitting isolated spectra lines with the same functional form presented in Eq. (5). The table with the number of coefficients used to describe each of the ILS parameters is presented here. The number of coefficients in the table corresponds to the degree of the polynomial (1 for constant, 2 for linear, etc...) used to describe each parameter in a specific order.

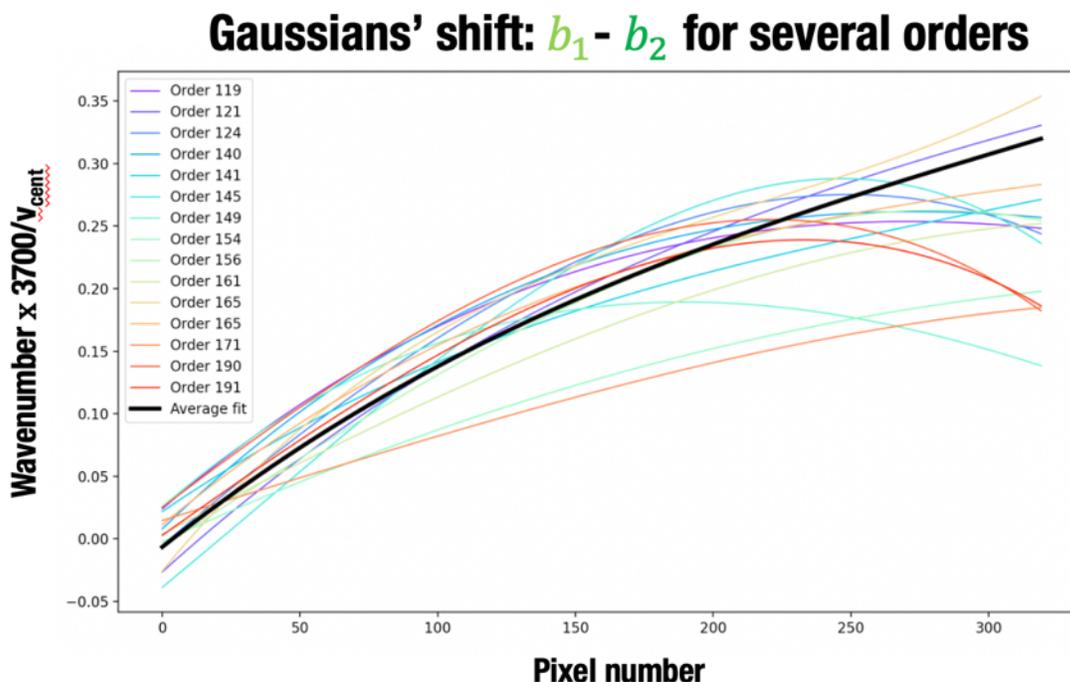


Figure 7. Polynomial fitting of the displacement between the two gaussians for each order, normalized to a common wavenumber. The average solution for all orders is in black.

$$\begin{cases} ILS_{1,i}(x) = A_{1,i} \exp \left[\frac{-(x - v_i + b_{1,i})^2}{2\sigma_{1,i}^2} \right], & i = \{0, \dots, 319\} \\ ILS_{2,i}(x) = A_{2,i} \exp \left[\frac{-(x - v_i + b_{2,i})^2}{2\sigma_{2,i}^2} \right], & i = \{0, \dots, 319\} \end{cases}$$

| | A_1 | A_2 | σ_1 | σ_2 | b_1 | b_2 | Total params |
|-------------|----------|----------|------------|------------|----------|---------------|--------------|
| BIRA | 2 | 2 | 2 | 2 | 2 | 2 | 12 |
| IAA | | 2 | | 2 | 2 | 2 | 8 |
| GSFC | | 1 | | 1 | | 3 or 4 | 6 |

Table 1. Number of parameters used in each parameterization, according to the latest formulation (Mar/2021).

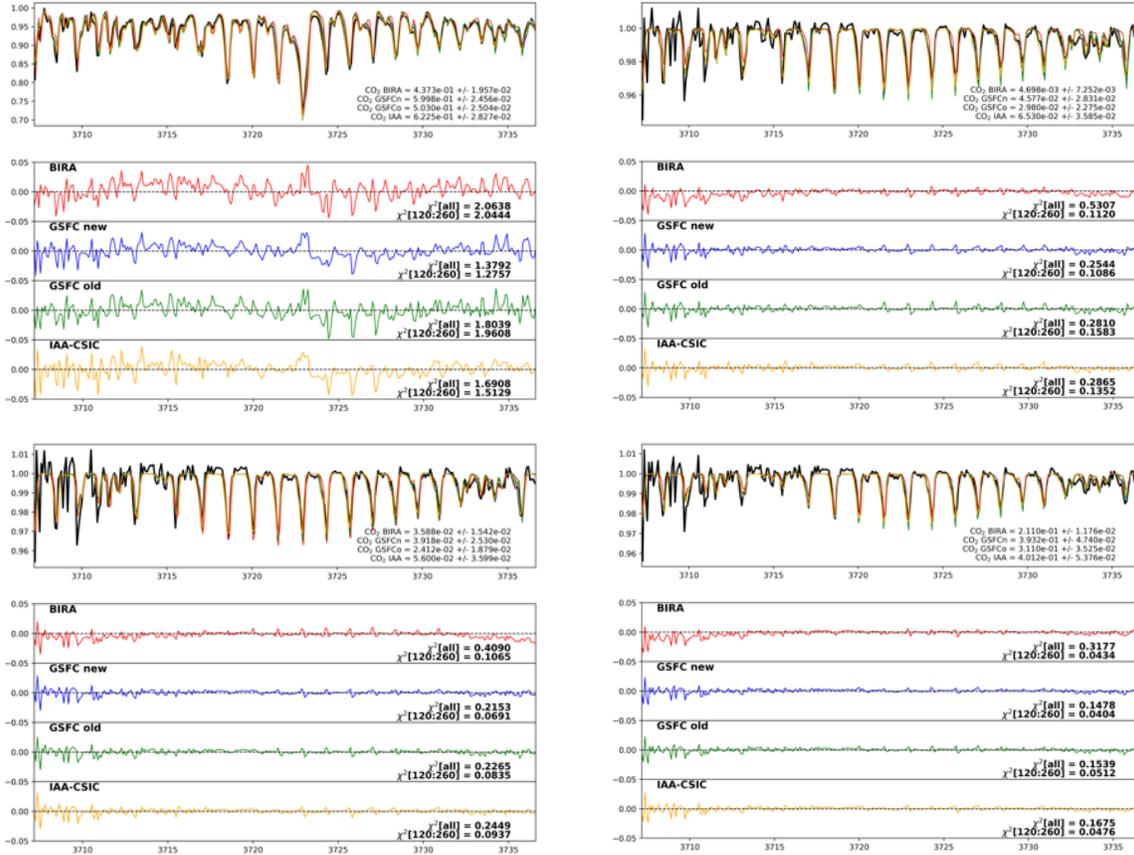


Figure 8. Comparison between retrievals performed using different ILS parameterizations. The “GSFC old” solution follows the same formalism as the new one, except for the degree of the polynomial of the displacement (second vs. third), and a different RP (20,000 vs. 17,000) and amplitude of the second gaussian (0.25 vs. 0.30). The altitudes, from top left to bottom right are 67 km, 103 km, 113 km and 141 km.

A comparison between these solutions has been done directly on actual retrievals: using PSG, we have generated high resolution radiances which include the latest AOTF parameterization by GSFC (Mar/2021). Then each ILS solution is used to generate NOMAD spectra, and a retrieval is performed at different altitudes (using a simple Levenberg-Marquardt scheme), to see how the solutions compare in terms of spectra residuals, which offer a relevant metric to test the accuracy of each ILS parametrization directly on data. The wavenumber calibration and the polynomial continuum are both iteratively corrected. The results of this comparison are presented in Figure 8 for different altitudes. All the schemes perform relatively similarly, which shows that a successful description of the ILS can be obtained with few parameters and susceptible to a high degree of simplification. Therefore, the search for a solution as much simple as possible and common to all groups is justified.

7. Impact of the ILS on retrievals: are A_2 and σ uniform?

One of the most significant limitations of the self-convolution technique in this particular case is that the actual estimation of the properties of the second gaussian can be problematic, mainly because of two facts: 1) the second gaussian is superimposed to the first one in a large portion of the spectrum, and 2) the self-convolution uses *a-fortiori* only a portion of the spectrum, rather than its local properties, making it difficult to characterize the punctual properties of the fainter, second gaussian.

| Order | Full spectrum A_2 | Full RP | Left spec. A_2 | Left RP | Right spec. A_2 | Right RP |
|--------------------------|---------------------|-----------------|------------------|---------------|-------------------|-----------------|
| 116 | 0.275 | 17,500 | 0.35 | 16,800 | 0.2 | 16,800 |
| 118 | 0.3 | 16,800 | 0.3 | 18,200 | 0.25 | 16,800 |
| 121 | 0.3 | 16,800 | 0.35 | 17,500 | 0.2 | 17,000 |
| 122 | --- | --- | 0.35 | 17,500 | 0.2 | 16,800 |
| 140 | 0.275 | 16,800 | 0.3 | 17,500 | 0.25 | 16,800 |
| 141 | 0.275 | 16,800 | 0.35 | 17,500 | 0.25 | 16,800 |
| 148 | 0.25 | 16,800 | 0.25 | 18,200 | 0.2 | 18,200 |
| 149 | 0.3 | 16,800 | 0.35 | 18,200 | 0.2 | 18,200 |
| 156 | 0.275 | 16,800 | 0.35 | 17,500 | 0.25 | 16,800 |
| 158 | 0.25 | 16,800 | 0.25 | 16,800 | 0.25 | 16,800 |
| 165 | 0.25 | 16,800 | 0.35 | 17,500 | 0.2 | 17,500 |
| 171 | 0.3 | 16,800 | 0.35 | 16,800 | 0.2 | 17,500 |
| Average | 0.277 | 16,863.6 | 0.325 | 17,500 | 0.221 | 17,166.7 |
| May/2020 Solution | 0.25 | 20,000 | 0.25 | 20,000 | 0.25 | 20,000 |

Table 2. Results from the analysis of each order, reporting the optimal (i.e., lowest RMSE) values for the RP and the amplitude of the second gaussian. In red the values whose improvement in RMSE is significant (>20%).

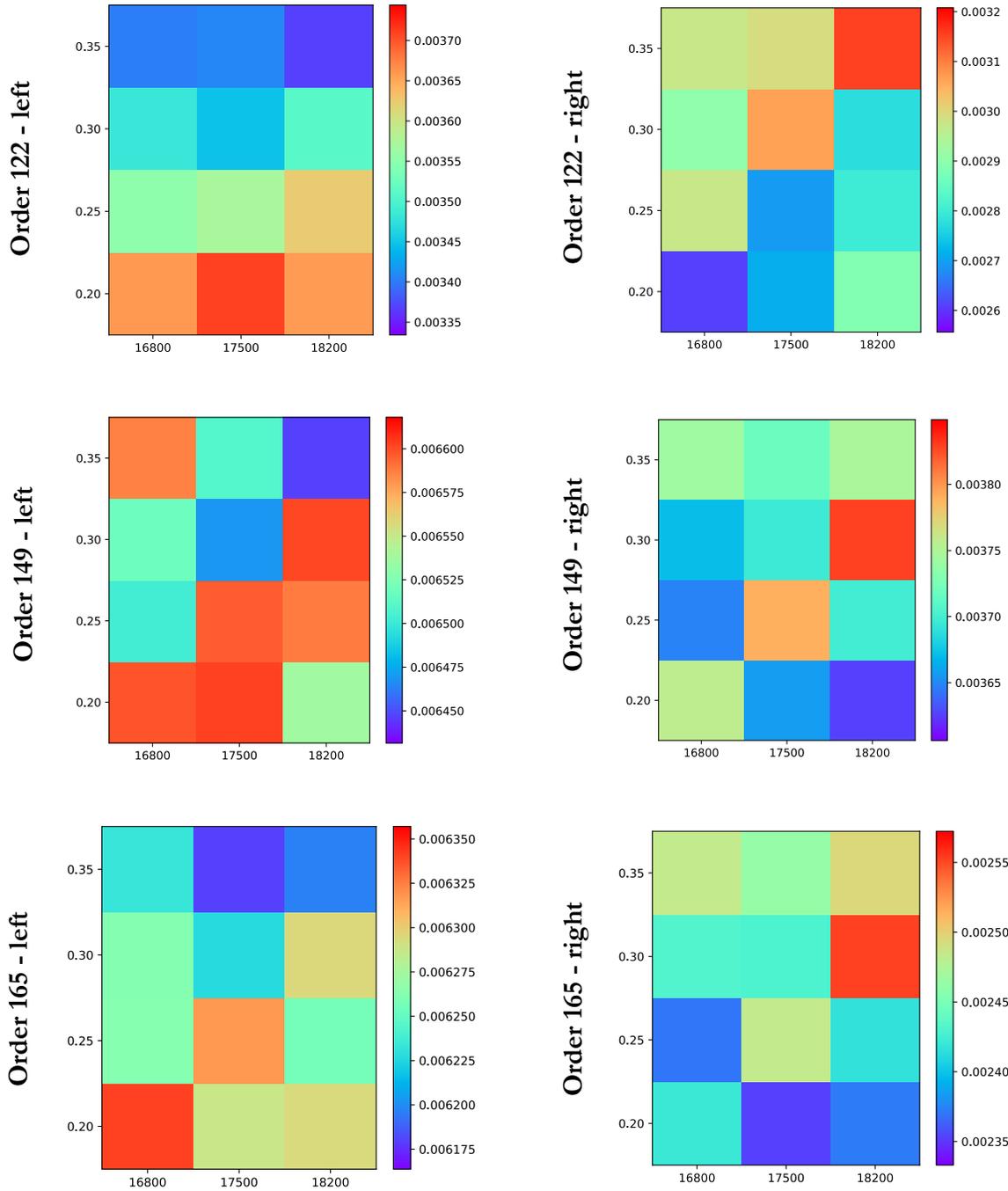


Figure 9. Sample results of the RP vs. A_2 analysis. The x-axis reports the RP, the y-axis the A_2 . The colorbar shows the RMSE.

A more sensitive way to test these properties is to work on actual retrievals, by iteratively changing these two properties and seeing what the real impact on the residuals and the retrieved parameters is. To this aim, we have performed several retrievals on the following orders: 116, 118, 121, 122 (altitude 35 km, no saturation); 140, 141, 148, 149 (42 to 50 km); 156, 158 (77 km); 165 (131 km); 168, 171 (40 km); for each spectrum, we perform a retrieval with RP of 16,800, 17,500, and 18,200, while the gaussian amplitude is 0.20, 0.25, 0.30 or 0.35, for a total of 12 retrievals per case. To test the accuracy, we simply look at the RMSE of each retrieval, and compile matrices of the RMSE values and see if any trend arises. This is done on both the left side and the right side of the spectrum as well, to identify strong trends within the same spectrum with pixel.

The results of this analysis are shown in Figure 9 (for some sample cases) and Table 2 (for all of them). We can see that there is no specific trend for the RP, neither across orders nor in the same order, while the \mathcal{A}_2 could have some variability across pixels, as the minimum RMSE in the left part of the spectrum is obtained in some cases with higher \mathcal{A}_2 values, and for lower \mathcal{A}_2 on the right. Nevertheless, the change of RMSE is significant ($> 20\%$) only in a minority of cases, therefore we chose to not to introduce any definite variability of neither the RP nor \mathcal{A}_2 in our final formulation of the ILS. As for the RP, indeed, this is justified by the explanation of the double gaussian as given in previous sections (i.e., “there is only one grating”).

The final recipe for the ILS is given at the end of this document, together with the procedure to create a synthetic spectrum.

8. Retrieval of full CO₂ profile with the new ILS solution

The new ILS solution has been further tested to verify its robustness throughout the whole NOMAD spectral range in the context of CO₂ vertical profile retrievals. This is an important diagnostic, since CO₂ lines exhibit a variety of degrees of saturation in different orders, therefore they can probe different altitudes and put to stress the ILS model.

We have chosen the same full scan used for ILS calibration purposes to retrieve the CO₂ column in a set of 16 diffraction orders, divided in 4 sets: 116, 118, 121, 122 (low-altitude); 140, 141, 148, 149 (mid-altitude); 156, 157, 158, 159 (mid-high altitude); 163, 164, 165, 166 (high-altitude). This choice is a compromise between the need to cover the whole atmosphere up to 150 km, and the fact that selecting nearby orders is very useful to verify whether this ILS solution gives consistent results, net of other effects that are excluded from the fit (in particular, temperature dependencies of high- j lines vs. low- j lines). The choice of the full-scan grants that the instrument is looking at the same atmosphere (net of lat-lon variability across altitudes) in all the orders taken at similar altitudes, given the acquisition time of a full orders cycle. For these retrievals, we assume an a-priori atmosphere as given by co-located interpolated MCD v.5.3 profiles, and we fit only a scaling factor to the a-priori CO₂ column along the line of sight, with no inversion of the overall profile.

The objective of this exercise is also to verify the presence of an offset, i.e., extra featureless flux that is not accounted for in the current modeling of the instrument. For a chosen column of a gas, the presence of an offset would result in a decreased line depth in the model spectrum compared to the case with no offset; when performing a retrieval, that would impact the estimated CO₂ column and generally result in an increase of the retrieved CO₂ column as the offset increases. It is expected also that the impact should be larger for saturated lines, as a small change in the apparent line depth results, in the curve of growth correspond to large changes in the column.

The way in which the offset δy is introduced into the simulated radiances is shown by the following equation:

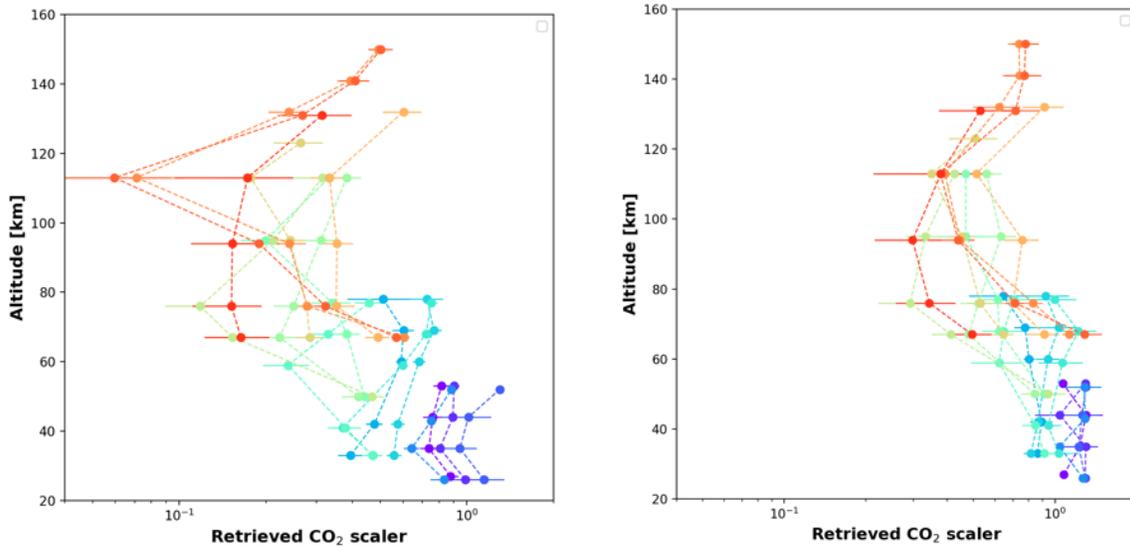


Figure 10. Retrieval of CO₂ profiles for different orders, without applying any offset (left) and with a 25% offset common to all the orders (right).

$$y_{off}(x_i) = (1 - \delta y) \cdot y(x_i) + \delta y \quad (7)$$

with the notation that follows previous equations. This is a way to account for the offset independently of its origin, which will be discussed later.

To explore the consistency of the ILS and the effect of introducing an offset, we have performed two retrievals of full profiles, without any offset and with an offset of 0.25 for all the orders. This second attempt is equivalent to say that 25% of the total observed flux does not come from the observed diffraction order and is featureless. The results of this test are shown in Figure 10.

Overall, profiles from nearby orders show a good degree of consistency among each other, with differences that are likely due to unaccounted (in the fit) temperature dependencies. This is the case of orders 163-166 vs. 164-165 (wing vs. center of the vibrational band) and of order 149 vs. 140-141-148. The most prominent effect, however, is the extremely low scaler retrieved for orders in the group 163-166 around 110 km. This is the altitude where these lines start to transition from a saturated to a non-saturated regime. When adding the offset, the effect is maximum here as expected, raising the retrieved scaler by a factor of 6. At lower altitudes, the offset has a much smaller effect on the retrieved CO₂. The agreement among groups of orders with altitude is acceptable within the a-posteriori uncertainties, however the idea is that the agreement can be further improved by assigning a specific offset to each order or group of nearby orders. Further analyses in this sense are needed to investigate the optimal offset for each order.

9. The origin of the offset

We have investigated the possible origin of a spectra offset of the intensity as deduced from this exercise, and we have isolated at least 3 possible explanations for it:

- 1) **Dark subtraction:** before the normalization with respect to the out-of-atmosphere radiance, spectra are dark-subtracted. If the dark used for the subtraction exhibits temporal variations across an occultation, this may result in a residual offset in the normalized spectra. However,

for NOMAD SO the dark has usually an intensity between 10% and 20% of the signal, and the variations of the dark are usually not higher than 10% of the dark. This results in the possibility that unaccounted dark could yield a 1% to 2% baseline, much lower than the effects observed on spectra from various orders.

- 2) **IIS functional form:** as said, in principle the IIS of a grating results from diffraction effects, and ultimately is described by a sinc function. When using a gaussian, the contribution from the wings is not accounted for, resulting in an offset. The effect would be exactly equivalent to the one simulated in the previous exercise, yet a 25% offset is very significant.
- 3) **AOTF offset:** the AOTF could be characterized by a baseline that introduces extra flux from the nearby orders. On a scale of ± 7 orders, a 25% baseline would result in an AOTF baseline of 0.016, or by a Gaussian base whose amplitude is scaled accordingly and whose contribution tends to 0 as we go further from the main order. This effect should be further explored, also because an AOTF baseline can change the depth of lines coming from different orders in a different way, potentially being easier to identify than a simple baseline due to the IIS.

In Figure 11 we show the difference between adding a simple 25% offset to the spectrum as done in the previous exercise and adding an AOTF baseline whose area over ± 7 orders is equivalent to that. The addition of the AOTF baseline has similar effects on the retrieved CO_2 column than adding a simple offset, with the further advantage of improving spectral residuals.

10. Final considerations: ILS recipe and offset recommendations

The recipe for the ILS as derived in this work can be described with 6 coefficients (not accounting for A_i), that are the following according to the formalism of Eq. (6):

$$\left\{ \begin{array}{l} A_{1,i} = 1.0 \quad \forall i \\ A_{2,i} = 0.3 \quad \forall i \\ b_i = \text{poly}(\beta_{0,\dots,3}, i) * v/3700 \\ RP = 17,000 \quad \forall i \end{array} \right. \quad (8)$$

$$\beta_{0,\dots,3} = [3.528 \cdot 10^{-9}, -3.3977 \cdot 10^{-6}, 1.7475 \cdot 10^{-3}, -6.4424 \cdot 10^{-3}]$$

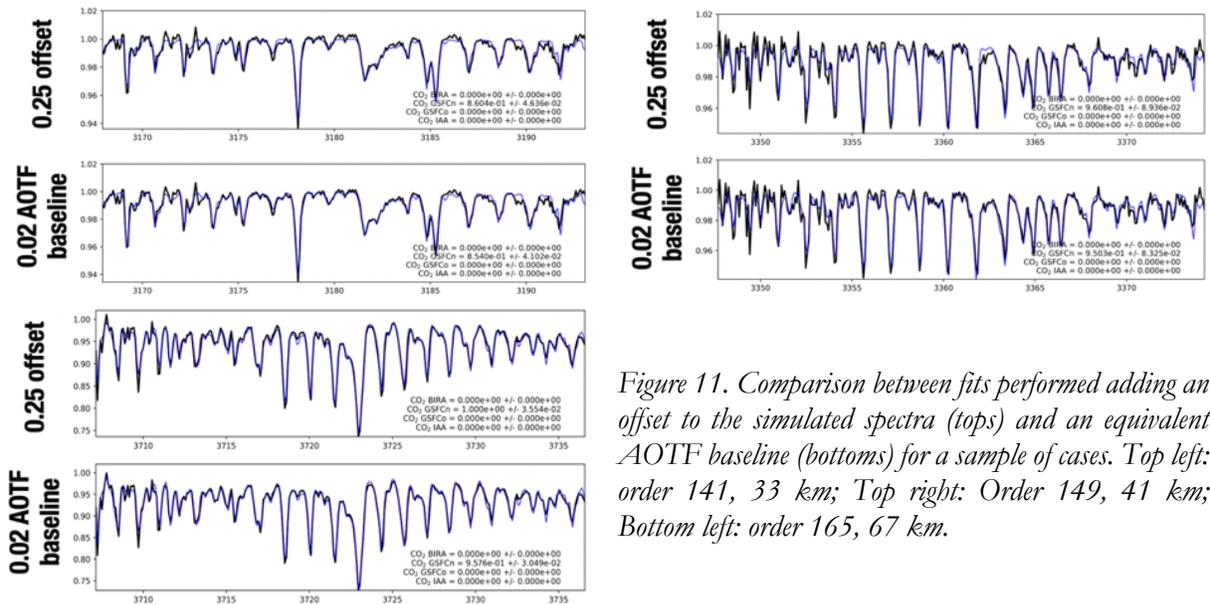


Figure 11. Comparison between fits performed adding an offset to the simulated spectra (tops) and an equivalent AOTF baseline (bottoms) for a sample of cases. Top left: order 141, 33 km; Top right: Order 149, 41 km; Bottom left: order 165, 67 km.

The simulation of a synthetic spectrum is done as follows:

- 1) A spectrum is synthesized at a resolution at least at 10x the NOMAD SO resolution, on a spectral range that covers the number of desired nearby orders.
- 2) Orders are summed according to the AOTF and blaze model.
- 3) Spectrum is convolved to the RP.
- 4) A second spectrum is interpolated with the frequency solution displaced in each pixel by b . This spectrum is multiplied by A_2 . This permits that a single standard convolution can be applied, and no special convolution/windowing approach is needed.
- 5) The two spectra are summed and divided by A_1+A_2 .

Based on the analysis presented in this report, the possibilities to refine this recipe would be the following:

- 1) Add another coefficient to describe a linear dependency of A_2 with pixel number.
- 2) Potentially simplify to a second-order polynomial the shift.

As for the offset, we will keep exploring the possibility of adding a small offset to the AOTF and evaluate the impact on the retrieved CO₂ columns at various altitudes.