

HowTo minimise fringe residuals in SWS data

Fred Lahuis, Do Kester and Dieter Lutz
freddy@srn.rug.nl

Abstract

This HowTo describes how to minimise fringe residuals in SWS data. In all SWS AOT-bands fringes are present to some extent. The minimisation of fringes in OSIA is handled in two steps. The first step is to reduce the amplitude of the fringe residuals when applying the Relative Spectral Response Function (RSRF). The second step is to remove these residuals as much as possible.

1 Introduction

Fringes originate on plane parallel surfaces in the light path of the instrument. These surfaces act as Fabry-Pérot etalons each of which can add unique fringe components to the source signal. In the infrared wavelength range surfaces separated by a few mm up to a few cm form the most efficient FP's. It is therefore not surprising that fringes are seen at most wavelengths in the SWS instrument. The fringes originate from a number of sources in the instrument. These can be filters, plane mirrors or detector surfaces. The last is the most conspicuous source of fringing for the band 3 detectors which utilise BIBIB detectors.

While in principle the step `RESPCAL` of the pipeline is designed to remove RSRF features such as fringes, there are several reasons why in practice fringe residuals are left after applying this step using the standard `CAL25` RSRF.

- Since the SWS spectral resolving power depends on source size, so does the RSRF. Point sources produce sharper and higher amplitude fringes than extended sources. The current RSRF is still fundamentally based on laboratory extended source data.
- For point-like sources, pointing shifts in dispersion direction and, at least for band 3, shifts orthogonal to the dispersion, change the RSRF and cause residual fringes. The same also applies to small wavelength calibration drifts.
- SWS01 introduces additional smoothing reducing the amplitude of the fringes.

For these reasons it will never be possible to derive a complete set of responsivity functions suitable for all observations. Therefore in OSIA routines are available to help minimise the fringe residuals in your observation as much as possible.

In most SWS AOT-bands fringes have been detected. In some cases they have only been seen in the RSRF at a very low level ($< 1\%$) in other cases they can only be detected in high resolution and high signal to noise observations and not in the lower resolution RSRF. Table 1 gives a quite complete overview of all the main fringe components and the amplitude

and type of observation in which they have been detected. In the paper by Do Kester [1] a complete description of the fringes in SWS data and fringe models for these are described.

band	D mm	range of D mm	RSRF %	Residual %	AOT
1A	4.20	4.1-4.3	-	8	6,2
1B	4.30	4.2-4.4	-	3	6,2
1D	4.45	4.3-4.5	2	5	6,2
1E	4.58	4.4-4.6	-	1	6,2
2A	4.43	4.4-4.6	1	2	6,2
	0.89	0.8-0.9	0.5	-	
	2.39	2.3-2.4	0.3	-	
2B	4.50	4.4-4.7	2	1	6,2
	0.89	0.8-0.9	0.6	-	
2C	5.86	5.6-6.1	4	4	all
	0.89	0.8-0.9	2	-	
3A	3.37+3.25	3.2-3.4	25	5	all
	5.26+5.72	5.0-6.0	10	5	
3C	3.25+3.38	3.2-3.4	17	4	all
3D	3.25	3.2-3.4	17	2	all
3E	3.28	3.2-3.4	19	2	all
4	13.25	13.2-13.3	0.2	-	-
4A	5.27+5.73	5.0-6.0	10	3	all
4C	6.67	6.5-6.8	1	-	all
4D	-	-	-	-	-

Table 1: Main fringe components per AOT band. The second column gives the effective thickness for the main components. Sometimes there is more than one component. These are listed separately when the components are quite distinct. Column 4 and 5 list the compound amplitude in percent, of the fringe components in resp. the RSRF and of the residual fringes in some observations on point sources. The last column list the AOT numbers where these residual fringes might be found.

2 OSIA modules

In OSIA three main routines have been developed to help reduce and remove fringe residuals. One of these (`RESP_INTER`) replaces the pipeline RSRF correction step (`RESPCAL`). The other two (`FRINGES` and `AARFRINGE`) remove the remaining fringe residuals after RSRF correction.

RESP_INTER replaces the pipeline `RESPCAL` step and tries to tackle the problem by countering the basic reasons: the RSRF is shifted and smoothed/enhanced (see section 3).

FRINGES works on an SPD or AAR (standard or rebinned) **after** an RSRF correction with `RESPCAL` or `RESP_INTER` has been applied. It implements a robust method of iteratively fitting sine functions to the data and correcting for them (see section 4).

AARFRINGE works at the level of rebinned AAR's. This is not optimal and whenever possible the above routines should be used. In cases where these fail it may be possible to achieve good results using `AARFRINGE` (see section 5).

The user should always be aware of the fact that the defringing of SWS data remains an interactive process which should be done with care requiring the judgement and common sense of the user. If applied wrongly, all mentioned methods and algorithms will introduce artifacts into the data which can lead to an erroneous scientific interpretation.

'Good data' for which excellent results are obtained typically consist of much clean continuum plus some lines. Even then, a rule of thumb may be that defringing rarely reveals complete miracles - lines and features which after defringing are good enough for scientific analysis usually have been suggested to the educated eye already in the fringing data by irregularities in the fringe pattern.

More complex cases are data with molecular bands covering most of the data, in particular if their frequencies are commensurate with fringes. Other problems occur for SWS02 data with just 1 or 2 fringes and broadish lines that blend into the fringes.

In most cases, doubt or no doubt, it is strongly suggested to try different reductions to assess the robustness and credibility of the results. This may be trying different methods or modifying the reduction parameters to the routines. Some examples are given in the sections below, see for example section [3.2.2](#), [4.2.1](#) and [4.2.2](#).

3 RESP_INTER

RESP_INTER is written to replace RESPCAL. It will reduce the amplitude of the fringe residuals through modification of the RSRF by matching the fringes in the RSRF with the observed fringes. It will shift and smooth or enhance the RSRF. It shifts the RSRF to compensate for wavelength shifts due to pointing offsets, source morphology or calibration drifts. The smoothing or enhancing is done to decrease or increase the amplitude of the RSRF fringes to compensate for the difference in resolution between the RSRF and the observational data.

RESP_INTER must be run after FLUXCON and before VELCOR. FLUXCON must be run first because the default is to derive single correction parameters for all detectors together. It is therefore important that the detector signals are flatfielded first to guarantee the most optimum result. This is different from the pipeline implementation where RESPCAL is executed first. Since the responsivity calibration and the absolute flux calibration are orthogonal operations their order can be reversed without any consequences.

In most cases there will still be small fringe residuals left after using RESP_INTER. This means that after RESP_INTER, extra processing using FRINGES is required. However, this does not mean that it can just as well be omitted. Experience has shown that the combination of the two routines in general gives best results. With RESP_INTER the remaining fringe residuals are better defined mostly because of the correction for wavelength shifts. Because of this FRINGES is able to find fringe residuals down to a much lower level.

3.1 Default operation

The defaults of RESP_INTER are chosen such that it is expected to produce good results for most data. By default RESP_INTER will:

1. Only modify the RSRF for bands 3 and off-band 4 data.
2. Smooth the RSRF for AOT-1 speed 1 and 2.
3. Enhance the fringe amplitude in the RSRF for AOT-1 speed 3 and 4 and for AOT-2 and 6.
4. Determine single correction parameters for the combined detectors. This gives stable results in all cases. For good signal to noise data a correction per detector may improve the results. ... try it and compare the results.
5. Fit the same continuum to RSRF and data. This means the same type of continuum fit over the same wavelength range.
6. Mask strong unresolved spectral features.
7. Correct per line and scan direction.

When the user does not specify any AOT bands all bands will be RSRF corrected. Only for all band 3 and the off-band 4A data RSRF modifications will be determined and applied. For all other bands RESPCAL corrected data is returned. When the user specifies particular AOT bands he should be aware that only those AOT bands will be RSRF corrected. This is to facilitate the possibility to make successive calls to RESP_INTER with different options per AOT band.

RESP_INTER plots the corrections and presents these to the user for inspection and approval. If the user is confident enough in the processing and knows his data to be without spectral features which are not masked properly he can tell RESP_INTER to continue QUIET-ly or he can process the observation with INTERACTIVE=0 and the /NO PLOT options. In this mode RESP_INTER can be used without problems for batch processing of e.g. large data sets.

3.2 Reduction examples

Below we give some examples of the way in which RESP_INTER can be used. For the explanation of the syntax you should consult the OSIA dc file.

3.2.1 Standard calls

These are few different ways in which RESP_INTER can be used.

```
spd = resp_inter( spd ) ; correct all AOT bands
spd = resp_inter( spd, aot_band = [ '3a', '3c' ] ) ; correct only bands 3A and 3C
spd = resp_inter( spd, aot_band = '3a', det=indgen(12)+25 ) ; correct per
detector
spd = resp_inter( spd, param=params ) ; all bands, return parameters for inspection
```

3.2.2 Spectral confusion

In some cases the major part of an AOT band is dominated by real spectral features, e.g. strong molecular bands of C₂H₂, HCN or CO₂ in band 3A. In that case it may be possible to derive a correction over a wavelength range not dominated by spectral features. Figure 1 shows an example where a single correction was determined between 15 and 16 μm and applied to the whole band.

```
spd = resp_inter( spd, aot_band = '3a', xrange = [15,16], nbin = 1 )
```

In this case it leads to a visibly better result reducing the amplitude of the fringe residuals by more than 50 %. This makes it possible to better constrain temperature and density profiles in the physical-chemical models and improve the detection of or fits to weak spectral features.

3.2.3 Using parameter input

If in a case like the one described in section 3.2.2 there is no useful wavelength range to be found it is possible to give correction parameters on input. The simplest is to use a single fringe amplitude correction.

```
spd = resp_inter( spd, aot_band = '3a', enhance = 1.2 )
```

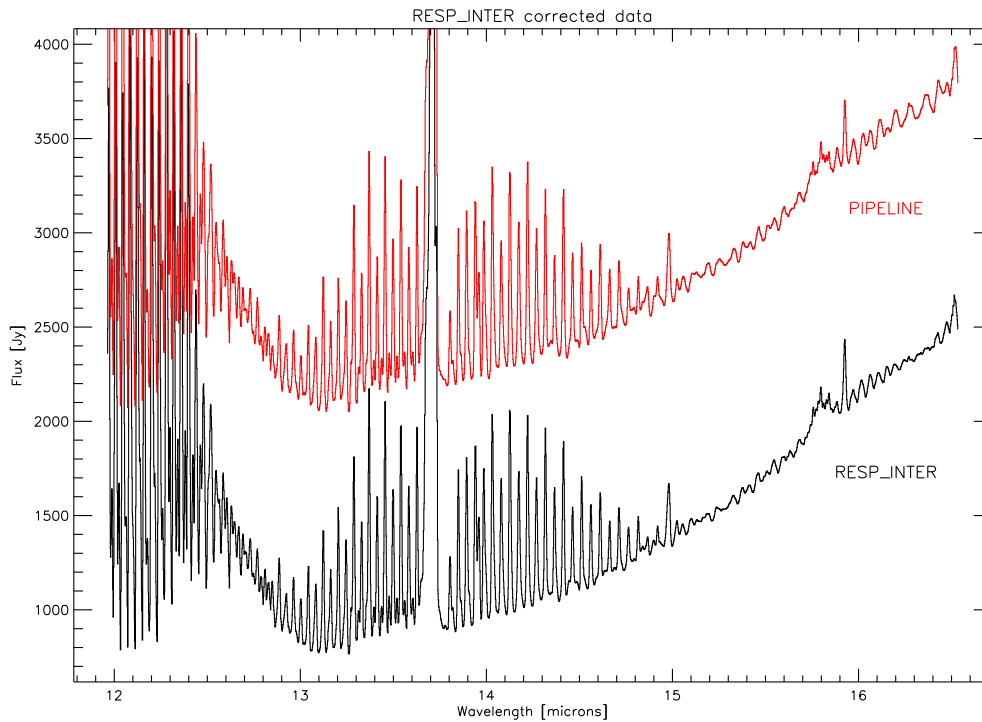


Figure 1: Band 3A data with dominant emission from several molecular bands complicating the correction of RESP_INTER and prohibiting a defringing using FRINGES. In this case a single correction is determined on the stretch of data between 15 and 16 μm which is not completely dominated by molecular emission. The result is a reduction of the fringe residuals by more than 50 % which leads to improved model fitting and better detection of weak spectral features. In black the RESP_INTER corrected data is plotted and in red the standard processed data using RESPCAL. The improvements can best be seen in the better underlying continuum of the C_2H_2 emission band between 13 and 15 μm and in the better defined C_4H_2 molecular emission band around 16 μm .

This will enhance the fringe amplitude in the RSRF for band 3a with a factor of 1.2. This is approximately the difference between the fringe amplitude of a fully extended source (RSRF) and a point source (most astronomical sources) for an AOT-2 or 6. Scientific analysis, e.g. improved and more consistent modeling of the data, should then prove the validity of this correction.

Another possibility is to use a source of similar extent to derive amplitude correction parameters and use these.

```
spd1 = resp_inter(spd1, aot_band='3a', /noshift, param=params )
spd2 = resp_inter(spd2, aot_band='3a', param=params, /use_params)
```

In this `spd1` is the reference `spd` and `spd2` the problem `spd`. `/NOSHIFT` is used since the shift is more observation dependent than the fringe amplitude. A small difference in pointing can result in significant differences in wavelength shifts (though small in terms of wavelength accuracy, ~ 0.1 resolution element). So to avoid making things worse it is better to leave the shift correction out.

4 FRINGES

The OSIA routine `FRINGES` models the residual fringes in the data after application of the RSRF by either `RESPCAL` or `RESP_INTER`. The latter is preferred since in general the results after `FRINGES` are then better than when using `RESPCAL`. This is a result of the fact that the fringe residuals are better defined after the corrections in `RESP_INTER`.

`FRINGES` approximates each residual fringe component with a sinusoidal function giving a total fringe model (see Kester [1] for the full details) I defined as:

$$I = \sum_{n=1}^N A_n \cos(4\pi D_n K) + B_n \sin(4\pi D_n K)$$

With D_n is the effective thickness of the etalon. A_n , B_n and D_n are the parameters to be estimated and K is the frequency in wavenumbers.

The basic steps involved in the processing in `FRINGES` are:

1. fit a smooth background and subtract this
2. mask out unresolved spectral features
3. fit N fringe components
4. make total fringe model
5. subtract the total fringe model
6. add background in again

The number of fringe components is determined either based on the calculated evidence or on the tolerance and number of fringes as defined by the user. The standard automated mode is using the calculated evidence. In this mode the procedure accepts a found fringe component and continues searching for a new component as long as the evidence is increasing. This is the advised mode for general use.

Figure 2 shows two defringing examples. The first is in band 1A where we see the fringes in the data for a point source but not in the (lower resolution) RSRF. What at first sight may appear as noise turns out to be mostly real fringes. The second

plot shows band 3A where the fringes are clearly seen in the RSRF and are still present in the RSRF corrected data. In both plots the input data, corrected data and RSRF are shown in the top panel and in the lower panel the background subtracted data and the composite fringe model is shown.

In good quality data fringe residuals can be detected down to a level of 0.1 %.

4.1 Advantages over FFT

There are three main reasons why the procedure adopted in FRINGES is superior to the use of FFT's.

1. The wavelength of the fringe can be estimated much better than in an FFT where it is limited to the Nyquist frequency, which in turn is determined by the number of data points.
2. It works on unevenly spaced data just as well as on a regularly gridded data set. Even when there are large gaps in the data like e.g. in an AOT2, it works fine.
3. It automatically determines how many fringe components there are present, avoiding possible overfitting or underfitting.

4.2 Using FRINGES

Over time FRINGES has evolved significantly. As much as possible the knowledge we have about the fringes and the pit-falls of the data have been caught in the routine and in the employed defaults. This means that in general the default use of FRINGES is expected to give good and stable results.

FRINGES is called to run on separate bands (combining the signal of all detectors and all scans) or on individual detectors. The input to FRINGES can be either an SPD or an AAR. In most examples we therefore use the name AARSPD to indicate this.

Some standard processing examples:

```
aarspd = fringes( aarspd, band = '3a' ) ; all of band 3A
aarspd = fringes( aarspd, band = 3 ) ; all of band 3A, 3C, 3D and 3E
aarspd = fringes( aarspd, indgen(12)+25, aperlist = 1 ) ; individual detectors of
band 3A
aarspd = fringes( aarspd, band = '3a', /updown ) ; all of 3A but up and down scans
separate
```

The processing is in general best done per (AOT-)band. It appears that the fringe residuals of the different detectors in a band are so much similar that real gains are made by combining them.

This does of course not mean that there are no cases where a non-default use of FRINGES will not give better results. The user should certainly not be discouraged to try what may at first sight appear to be unconventional uses of the routine. This is usually the best way to learn something new about the problem.

Below we will try to give some examples of 'advanced' use of FRINGES.

4.2.1 Up and down scans

This is a test to decide which is the best defringing, a defringing with all data combined or with the up and down scans defringed separately.

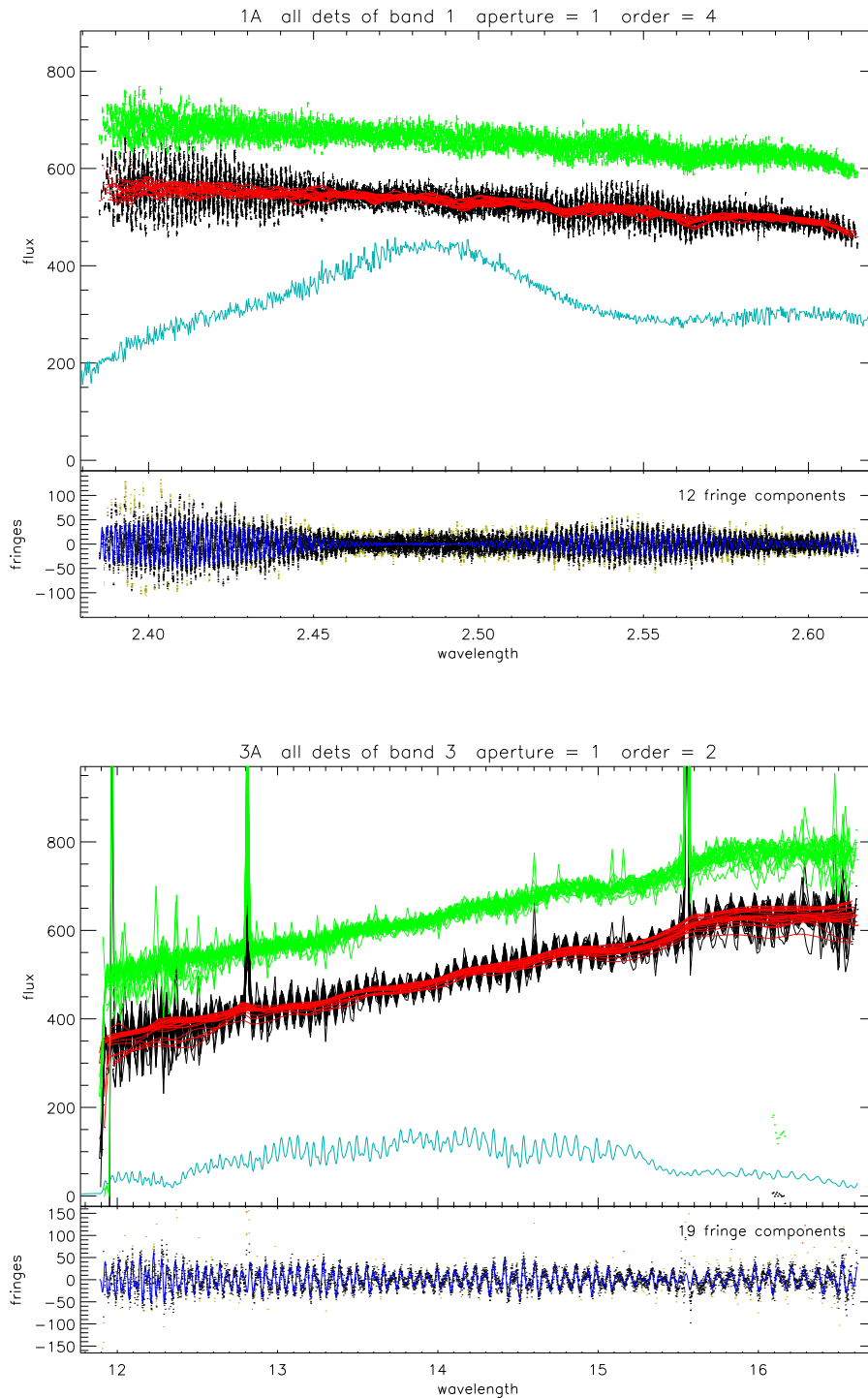


Figure 2: Defringing examples using FRINGES on band 1A and band 3A data. In the upper panel the data and the RSRF are shown. The input data is in black with the smooth background in red, the corrected data is plotted in green and is shifted for comparison and in light blue the RSRF (scaled to fit the plot). In the lower panel the background subtracted data is shown in black and the composite fringe model is overplotted in blue. Note that in band 1A none of the fringes are present in the RSRF while in band 3A the similarity between the fringes in the RSRF and the residual fringes is striking.


```
aarspd1 = fringes( aarspd, band = '3A' )  
aarspd2 = fringes( aarspd, band = '3A', /updown )
```

For this we use the evidence. The value in the last column of the IDL output of `FRINGES` 'LogPr' is the logarithm of the probability for that solution. When the up- and down-scans are reduced separately we may sum the probabilities. The difference of this with respect to the probability derived from the other processing gives us the evidence of one solution wrt. the other.

For the processing of the combined scans (to get `aarspd1`) we find a value of 2143.6. For the processing of the separate scans (to derive `aarspd2`) we find values of 1040.9 (UP) and 1027.2 (DOWN) adding up to 2068.1.

So the evidence of solution 1 with respect to solution 2 equals $10 * (\text{LogPr1} - \text{LogPr2})$ dB being 755 dB. We therefore decide that `aarspd1` is to be preferred over `aarspd2` by more than 750 dB. A significant difference. A visual check by overplotting both results will confirm this.

4.2.2 Limiting the cycle range

In the case when e.g. molecular bands are present which have a beat frequency of the same order as the fringes we can do the following.

1. Reduce the SPD data using the pipeline modules `ANTIMEM`, `DARK`, `FLUXCON`. This is then optimally reduced with `RESP_INTER` and `FRINGES`.
2. Find the optimum reduction parameters for `RESP_INTER` and `FRINGES` by repeating the following until you are happy with the result.
 - (a) run `RESP_INTER` with modified parameters, e.g. limited wavelength range or use less bins than normal, see section 3.2.2 for an example
 - (b) run `FRINGES` modifying e.g. the cycle range (see table 2) and wavelength range
 - (c) make a fit of the molecules to the reduced spectrum
 - (d) make a synthetic spectrum over the same wavelength range as the data
 - (e) judge the quality of the reduction by:
 - i. estimating the amplitude of the fringe residuals in the data after defringing
 - ii. defringing the `CAL25` to get a second estimate of the amplitude of the fringe residuals
 - iii. estimating the effect of the defringing on the synthetic spectra
 - A. defringed with `FRINGES`
 - B. molecular fit to defringed synthetic spectrum
 - iv. make a scientific judgement on the results
3. Make a final fit to the spectra and use the results from 2-e in the error estimation of the fitted molecular parameters.

This procedure has actually been used with quite some success in a few scientific programmes. One example of this can be found in Lahuis et al. [2] Figure 3 shows an example from this paper. In cases where the extent of the molecular band is too large this may not work and the best results will be the optimised `RESP_INTER` corrected data. An example of this is shown in section 3.2.2.

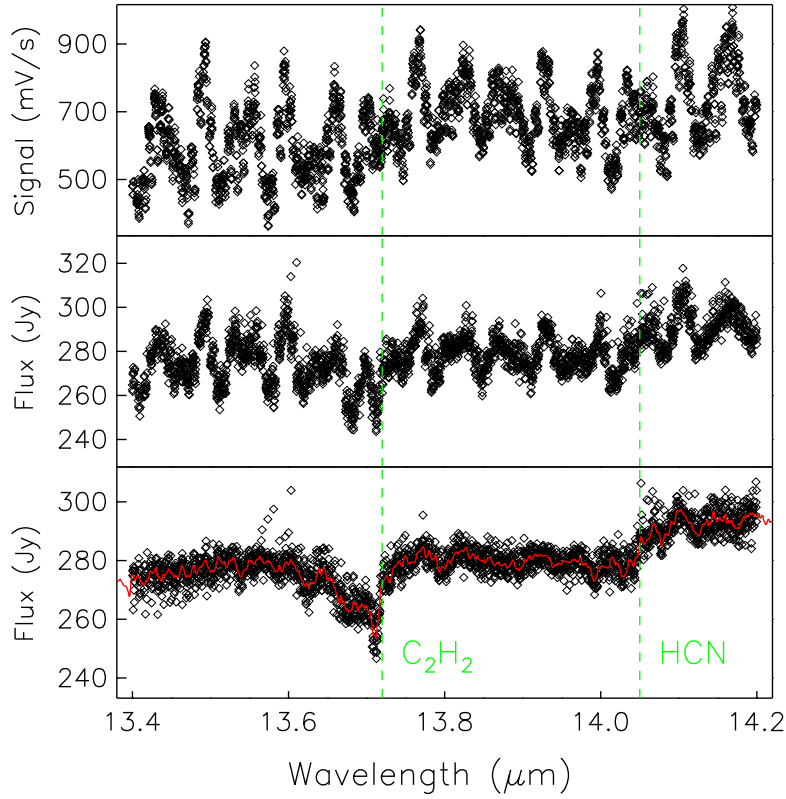


Figure 3: Example of optimal defringing as described in section 4.2.2. In this example we zoomed in on the wavelength range around 14 μm where we find gas-phase C_2H_2 and HCN . The wavelength range used in the defringing was larger, app. from 13 to 16 μm . In the top panel the raw AOT-6 SPD data from one detector is shown. In the middle panel the RSRF corrected data for all detectors and in the lower panel the defringed data very clearly showing the molecular absorption data.

band	aperture	cycle	ncycle
1	all	4000	1000
2	2	4000	1000
2	3	5500	1000
3	1	3000	3000
3	2,3	3000	600

Table 2: Default search range for fringe components for each band in FRINGES.

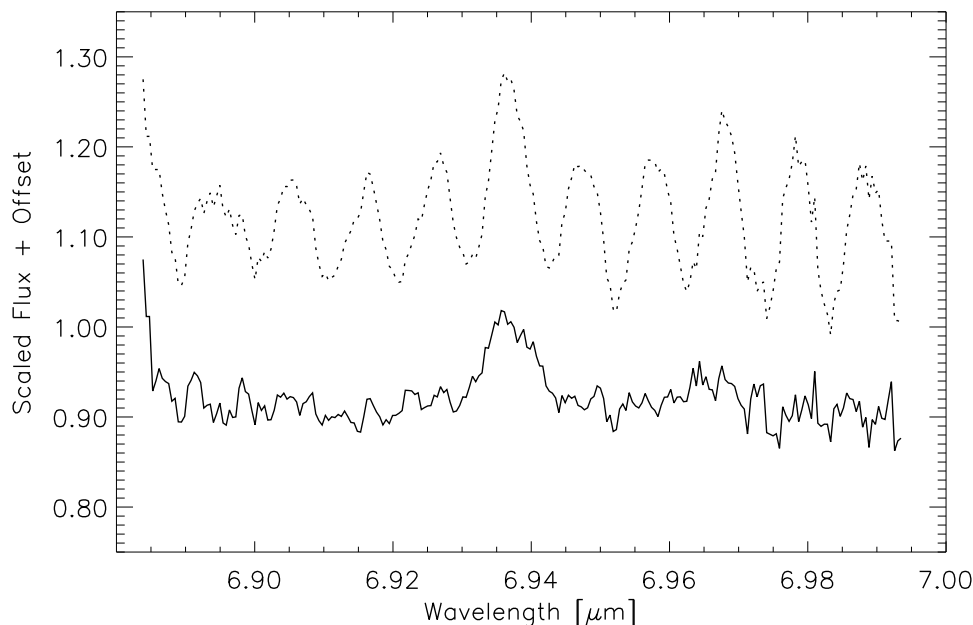


Figure 4: An example of a single SWS02 scan that can be defringed well using AARFRINGE. A relatively narrow line is superposed on a significant number of well-defined fringes.

5 AARFRINGE

The last routine available to tackle fringes is AARFRINGE. It is written to in principle only work on rebinned AAR data. It should be noted that it is advised to first try to do the defringing using RESP_INTER and FRINGES. In cases where these fail it may be possible to try defringing using AARFRINGE. The calling syntax is quite simply

```
aarout=aarfringe(aarin)
```

where `aarin` is the already processed and rebinned input aar. Details of use are described in the OSIA dc-file. The major steps to be taken within AARFRINGE are

- Selecting an AOT band. Fringes are coherent only within AOT bands, so that bands have to be treated separately.
- Masking out outliers that may occur at the edges of the rebinned spectrum, and strong lines. These might fool the fringe location algorithms. The DOK method (see below) is most robust to such problems but even there it is often useful to take out such regions before running the fringe detection.
- Selecting one of the three methods:

FFT This interpolates over peaks in the Fast Fourier Transform of the data to take out fringes. It works best with long stretches of data like entire AOT bands but often gives imperfect correction at the edges of this range.

DOK This implements the robust iterative sine fitting algorithm introduced by Do Kester for FRINGES, but working here on the rebinned AAR. Often, it provides the best result. NB. the AARFRINGE implementation of this method is out of date compared with the FRINGES implementation.

SIN This is a straightforward least-squares sine fit. Because of convergence problems, it is useful only for data with few fringes like single SWS02 scans where FRINGES cannot find proper fringes.

- For all methods, frequency ranges have to be defined in which fringes should be searched for. The OSIA dc file gives more details on this as well as a table which provides some hints concerning known fringe frequencies ranges for the various AOT bands.

5.1 AOT-2 scans

In some cases when we e.g. have only a single AOT-2 scan within an AOT band FRINGES does not work optimally because the baseline is just too short. In that case it is possible to achieve good result with the sine fit option in AARFRINGE. Figure 4 shows an example of such a case.

6 Document change log

Version	Date	Initials	Comment
1.0	19 Sep 2001	FL	First draft based on old HowTo's and Do's legacy paper
1.1	01 Oct 2001	FL	comments from Do
1.2	15 Oct 2001	FL	comments from Dieter and Russ
1.3	04 Jan 2001	RH	proper bibliography
1.4	27 Oct 2003	AV	edits and updates OSIAv 4.0

References

- [1] D. Kester, D. A. Beintema, and D. Lutz. SWS Fringes and Models. In *The Calibration Legacy of the ISO Mission*, volume ESA-SP 481, page 375, 2003.
- [2] F. Lahuis and E. F. van Dishoeck. ISO-SWS spectroscopy of gas-phase C₂H₂ and HCN toward massive young stellar objects. *aa*, 355:699–712, March 2000.