How to reduce SWS AOT02 Data in the ISAP GUI

Eckhard Sturm

May 3, 2000

Before you begin:

This guide is intended to give a basic description of how to use ISAP in the GUI mode to reduce the Auto Analysis Results (AARs) for SWS AOT02 data: it explains the basic principles and shows in a representative way how and in which order to work through each of the possible steps of data processing. It assumes that you are working with ISAP version 2.0. For a full description and alternate uses of a particular routine, consult the ISAP help. This help is accessible on the web: http://www.ipac.caltech.edu/iso/isap/isap.html or http://www.mpe.mpg.de/www_ir/ISO/observer/isap/isap.html. From within the GUI simply use the "Help!" button and select "WWW pages." The help pages contain - among other things - a beginners' GUI guide, a description of all the GUI Buttons, and links to the detailed help text of each individual ISAP routine ("Command Line Routines"). Here you'll also find the necessary information about routines that are not mentioned in this recipe. A lot of fundamental information about the SWS instrument and its behaviour is contained in the SWS handbook. The latest version is always available on the web: http://www.iso.vilspa.esa.es (follow the link to Documentation). Before trying to reduce SWS data we **strongly recommend** to carefully read at least the beginners' GUI guide, which will make you familiar with the basic principles of the ISAP GUI, and the SWS handbook!

If your AAR has been produced with a pipeline version earlier than 7, you should seriously consider/ask for a reprocessing of your data. The pipeline version can be found in the header of the FITS file. Note that the data in the official ISO Data Archive are all processed with version 8 (or higher). Also note, OSIA and IA3 do not update the pipeline version in the header (but you can use the header editor in ISAP to overwrite it manually).

These instructions suggest saving the intermediate products of the data processing after particularly time-consuming steps; the user should weigh the value of additional saves against the constraints of disk space.

Before starting ISAP, you should close all applications which use color (e.g., Netscape, or additional IDL sessions), as they will interfere with the ISAP color table. After ISAP has been started, you may restart these applications without affecting the ISAP session.

Finally: Whenever you have comments, suggestion or complaints about this manuscript (e.g., if you are missing some information), please let me know! Please check the web site for modifications of this document - from time to time.

1) **getting started / data input**

Start ISAP in the GUI mode (possibly "isap" or "isap_gui" - see the readme text of your installation). The first thing to do of course is to read in the data: Click on "Read data..." and then select "FITS" (provided you have your AAR available in its original FITS format). Choose the desired file from the file selection tool which pops up (i.e. double click on the file name).

2) **Quick look - I**

To see the full data set, select all scans, detectors (or bands), lines, and scan directions, then click on plot. For a closer (but still quick) look at your data plot the spectrum line by line. Make use of the different color coding options offered by ISAP (click on "Select Plot Style"). Note, that **plotting** of large data sets in plot modes **other than the default** "dot cloud" mode can take a **long time**. If you want to interrupt a plotting process, and you are running UNIX, type "Ctrl-C" in the command line of your IDL window, and then type "return". You should now be able to reset the plot style and to plot again.

Here follow some criteria you can use for a first cut evaluation of the data quality and to spot potential problems:

- *Are the flux values non-negative (on average)?*
  If the answer is NO, then the dark current subtraction in the automatic pipeline processing has not been optimal (the dark current has been over-subtracted). In this case you could very likely improve the result by going back to OSIA (the Observer's SWS Interactive Analysis system, which operates on SPD level) and performing an interactive dark current subtraction (with OSIA's routine dark_inter). Note that *on average* means that many individual points might still be negative in the case of a faint, noisy source.

- *Are there inconsistent trends in the continuum shape?*
  If some of the detectors within a line have continuum slopes that are tilted with respect to the other detectors, this could again be due to problems in the automatic dark current correction. Going back to OSIA is recommended, if a sustantial part of your data is affected.

- *Are there significant systematic differences between the two scan directions?*
  This is hard to tell on the quick look level, if you have many scans of a line. But if you have the impression, or if spot checks (i.e. plots of single detectors and single scans) show these differences, it is recommended to use OSIA's tool updown_inter, which is able to remove such differences on a level before the relative spectral response function and absolute flux calibration are applied to the data (but see also step 7)

below.)

3) **Remove bad data - I**

Sometimes the pipeline AAR (prior to version 7) contains data with a scan counter 0, line number 0, and/or a scan direction 0. Those data usually can be dropped. Just don't select scan 0 and scan direction 0 before you plot the data in the following steps. Also, if there are any data of band 5 or 6 (the FP bands) in your SWS02 data, they can be disregarded, too. In this case select all scans, lines, scan directions and bands, except those with a zero value and except bands 5 and 6, plot everything (select the default plotmode, i.e. unconnected, no color preference, which is fastest), and catch everything with the right mouse button. Choose "make aar" in the tool box, enter a new name and store the result.

4) **Split**

It is recommended to process the individual lines separately. Click on "Special Fns..." and choose "Split Apart AAR". Then check the "lines" box and apply. This will automatically split the AAR according to the line tag, and will store the result in "Stored Data Sets..."

5) **Start loop**

Now you can select the first of the line spectra: click on "stored data sets...", then on the name of the first line (name=filename_l1) spectrum, and then on "make prime" (under "Apply function to"). The following steps describe what should be done for each of the line spectra separately (i.e. in a loop).

6) **Remove bad data - II**

The next step should be the removal of all kinds of bad data. You can choose between "Tippex", "Mask" and "Zap" (invoked via the tool box which pops up if you select some data with the right mouse button). With TIPPEX a single point is replaced by a linear interpolation between the two corresponding neighbouring points. This is done as unique as possible, i.e. the two points choosen as basis for the linear interpolation are of the same detector, scan and scan direction, if any of this information is available. MASK simply masks out the selected data, i.e. they will not be taken into account for the computation of correction functions in the subsequent processing steps (e.g. in AVERAGE they won't contribute to the average). But the masked data stay in the data set, and shift, for instance, shifts masked points as well (without using them for the computation of the correction). With ZAP you really get rid of the data, they are thrown away. To remove single outliers you could use TIPPEX, but at this stage it is usually a good idea to simply ZAP all bad data, because the redundancy in the data and the spectral sampling is high enough.

In a first step you should plot all the data of that line (i.e. all scans, dets, sdirs) in the default "dot cloud" plot mode (i.e. unconnected, no color code). You will probably see many clear outliers, which you can zap away right now. In a second step, you should have a detailed look at every single scan and every single detector. This will be the most time consuming step, but also one of the most important ones. You can significantly improve the quality of the final spectrum, at least for moderate to faint sources. The recommended plot mode is: connected, color code scan direction. Choose the first scan and go through all 12 detectors for that scan (make sure the band button is off!). Then choose the next scan etc. You can plot several detectors simultaneously. But when in doubt, plot detector by detector. Zap all clear outliers.

Besides single outlying points another kind of "bad data" can be sudden jumps in flux (a glitch) followed by a subsequent exponential-like decrease over several seconds/data points, due to memory effects. This effect is particularly significant in band 4. You should zap the entire memory tail.

In the decision of what is a glitch and what is a line you should make use of the redundancy in your data: If the feature is seen only in one of the twelve detectors it is a glitch. If it is seen in several or all twelve detectors at the same time but at slightly different wavelengths it is a glitch which caused some electronic cross-talk (that's why it is seen in more than one detector). A line is only a feature which is seen in at least most of the detectors at the same wavelength and which is present in both scan directions (up and down, or -1 and 1). The different possible color codings (per detector, per scan, per scan direction), adjustable via the "Set Plot Style" button, and the ability to plot different detectors, scans or scan directions separately, are usually a great help in this business.

E.g. you could select "connected, color code detectors", which helps identifying "bad" detectors: Sometimes a whole detector seems to produce much worse quality than the others (e.g. detectors 21, 26, 29 or 34 are known to behave strangely sometimes). This can be checked if you set the color preference to "Detector Number." With the MIDDLE MOUSE BUTTON you can always get information about a single data point. Just click on the desired point (or near to it). In the info field at the bottom of the GUI you'll get e.g. the detector number, scan number, scan direction etc. In this way you can identify the detector number (if you have plotted more than just one detector). To get rid of a whole detector use "Zap assoc", which will zap all points associated with the selected data points (,so, simply select a few points of both scan directions for the detector under consideration with the right mouse button and apply "zap assoc").

If you have to zap (mask/tippex/etc.) a lot you can make use of the "Preselect" button: Click on it and choose "Zap". From now on ISAP will automatically zap whenever you catch data points with the right mouse button. Don't forget to switch preselect off when you are done with zapping. (If you do forget, the "undo" button in the upper right corner of the GUI might come in handy...).

An alternative to zapping "by hand" is automatic "Clipping".

While going through the data detector by detector, and scan by scan, you will also realize whether your data show significant differences between the two scan directions (hence the plot mode set to "color code scan directions"). This topic is discussed next.

Don't forget to STORE!

7) **Differences between up and down scans**

Each AOT2 scan actually consists of two scans of different SCAN DIRECTION:
+1: up scan (increasing grating positions, decreasing wavelength)
-1: down scan (decreasing grating positions, increasing wavelength)
The sequence is always up →down. (Hence, each scan actually yields 24 spectra - 12 detectors x 2 scan directions.) Due to hysteresis/memory effects of the detectors there can be differences in the shape of the continuum between the up and down scan of a given scan number (see the SWS Handbook). In ISAP there is no optimal tool to correct for this effect. If these effects are strong, you should go back to OSIA and try to correct them on SPD level. In ISAP you can at least mask out or zap the portion (sometimes a full scan direction) which you believe to be the worse/more effected one. Usually the down scan is considered the more trustworthy scan, because it follows after the up scan in the observation sequence, i.e. the detectors have had more time to settle. However, this is true only if the memory effect was caused *before* the scan sequence started (by a large signal between the science scans, e.g. a photometric check), and if there was enough time for the down scan to settle. If the memory effect is caused *during* the scan itself, e.g. by scanning across a strong line, or by strong cosmic ray hits, both scan directions may be similarly affected. In this case an averaging of both scan directions might help. This is discussed in the following step.

Before you go on, it is certainly a good idea to STORE the result of this time consuming work. Consider perhaps even to write it to disk (select the data under "Stored Data Sets...", click on "Write to Disk in..." and choose the format).

8) **Average - I**

Now you can go on and average the two scan directions for each scan and detector. Select all data and choose "Average". For a detailed description of the average tool please read the web pages (see "GUI Buttons"). Average across scan direction (but not across detector or scan), i.e. don't select the scandir button (but do select the det and scan button).

The bin size should correspond to the spectral resolution. The resolution of an AOT2 observation can be found e.g. in the SWS Handbook. For a point source it is roughly in the range between 1200 and 3000 (resolution element (FWHM) = 0.015 to 0.004). Extended sources have ≈ 0.6 - 0.7 times lower resolution. (Note that the resolutions quoted in the handbook are for extended sources!). Measured resolutions can also be found in
http://isowww.estec.esa.nl/instr/SWS/doc/sws_ip/.
To preserve the oversampling the 12 detectors provide, the bin size should be smaller than the actual resolution element [1]. FWHM/5 is usually a value which gives good results.

[1] For SWS02 the step width can be understood as follows: the wavelength range of one detector covers appr. 10 resol elements, scanned in 50 steps. E.g. @ 10 μm: resol ≈ 2000 →resol element ≈ 10/2000 = 0.005 →step width ≈ 10 x 0.005 / 50 = 0.001 μm. For band 4 there are two scanner steps per reset interval, e.g.: @ 34 μm: resol ≈ 1500 →resol element = 34/1500 x 2 = 0.045 →step size ≈ 10 x 0.045 / 50 = 0.009 μm)

These FWHM/5 (or resol x 5) values are offered as default values in the average tool (whenever ISAP is able to automatically deduce the resolution element from your data, i.e. whenever the data are unique and the status word is present).

The plot style for the "before" window should be the default style (unconnected, white, dots). The plotstyle options in combination with the possibility to zoom in (with the left mouse button), e.g. into a small range around a line, and to freeze the plot range in both windows to identical values, should help you to judge the result. To adjust the plot ranges in both windows use "Replot Max" in the left window, then use "Copied Range" in the right window.

If the oplot button is checked, ISAP overplots the result directly in the left hand side plot window. At this stage of the processing the plot could get very crowded (in particular if you have many scans). Therefore, you should switch off the oplot mode. To check if your bin size is appropriate you could perform a few spot checks prior to the actual averaging. Just go into average 3 or 4 times, with only one detector and one scan at a time. Here you can switch on the oplot button. If you use the default value for the bin size (or resolution), you should perhaps write it down, because in subsequent calls of the average routine ISAP will not be able anymore to deduce such a default value from the data.

In oplot mode you will also see the points that have been clipped by the clipping routine (in red). Note, that the small bin sizes associated with sws spectra can be too small for a significant clipping. In this case try to lower the sigma value (e.g. 2 or 1). Alternatively, there are two stand alone clipping routines available. A very basic one is available under the "arithm" button (in the tool box). It performs a kapp x sigma clipping on the basis of all selected data (i.e. the bin size is the entire selected wavelength range). Kappa can be specified by the user (e.g. 2.5). Outliers more than kappa x sigma off will be masked. You can then use the plot option "show masked data" to view the result. If you like it: "zap masked", if not: "unmask". A more advanced clipping tool is available under "Clipping". This routine allows you to adjust a bin size.

You can choose between different methods of averaging (see the description on the web). Try all of them to find the best method for your data set.

Don't forget to STORE!

9) **Flatfielding (shift)**

Although the best effort has been made in the offline processing pipeline to ensure that the variation in response between detectors has been accounted for, there will inevitably be some residual discontinuity between detectors. To normalise all detector data to a common level use "Shift". You can either adjust all 12 detectors to the mean level or adjust 11 detectors to the level of a 12th, "good" detector. Please see the web help for more details. If the signal is weak the differences between the spectra of the 12 detectors can be attributed to uncertainties in the dark current subtraction. Hence, the "offset" mode should be choosen. Only for strong signals (above several 100 Jy) the dark current can be neglected and differences are due to gain effects, i.e. multiplicative effects. In this case "scaling" would be the right option. No general rules can be given here to decide which of the modes is the best. The same is true for"weighted" or "unweighted". However, in many cases "offset", "preserve mean", and "unweighted" is a good default.

An example for a problematic case is the combination of a strong line on a weak continuum: the continuum may be better corrected in "offset" mode, whereas the line itself might need a "gain" correction. If the continuum flux is around zero, both modes may fail, however. In this case you could try the following: mask out either the line or the continuum (depending on what you are interested in most) and shift in offset or gain mode. Then unmask again. (Note: shift will not use the masked data for the computation of the correction, but it will apply the correction to the masked data as well).

Plot your spectra scan by scan, with the plot mode set to "connected, color code detectors". Run shift scan by scan. You can also select several scans at once. Shift will then shift detectors scan by scan, but only in "preserve mean" mode (and it is harder for you to check the result).

For band 1 and 2 data, flatfielding usually does not improve the data quality significantly, but for band 3 and 4 data this can be a very important step.

Be aware, however, that shifting in offset mode, i.e. adding or subtracting a constant, at this stage of processing is in principal wrong, because the relative spectral response function has been applied to the data (multiplicative). Nevertheles a shifting at this stage can improve your data quality. The result of the shifting should in any case be examined critically.

10) **Remove bad data - III**

After shifting you might want to do some additional zapping.

11) **Shift across scans**

Plot all the data of the current line with color code set to "scans". If there are no significant offsets between the scans you can continue with step 12 and average across scans and detectors in one go. However, if there *are* differences of the average flux level between the scans, you might consider to first remove these differences before you average across them. In this case average across detectors, first, – see step 12, but with the scan button on. Then use shift to remove the offsets between the different scans (shift automatically shifts scans – not detectors – when you call it with a set of scans without detectors). Now you can go back to "Average" and average across scans. (Note: there is a difference between 1) shifting in offset mode and then averaging, and 2) just averaging: in the first case you preserve the mean of the entire spectra, in the second case you preserve the mean at each wavelength. Multiplicative shifting and subsequent averaging is obviously also different from just averaging.) For weak sources this additional shifting across scans can further improve a little bit the final S/N. Be very careful when using normalize.

12) **Averaging - II**

Now (i.e. you had no differences between scans) you can go on and average the spectra of all 12 detectors and all scans to get a single, mean spectrum. Select all data and choose "Average". Average across detectors and scans (buttons off). The bin size should again correspond to approximately 1/5 of the spectral resolution element.

This time the oplot option could be very useful in judging the result.

Don't forget to STORE!

13) **Defringe / cf. RSRF**

For strong sources you may see periodic modulations in the continuum of band 2 and 3 data (and sometimes band1, too). These are the fringes, which are described in more detail in the SWS IDUM. In an AOT2 you may, however, see just a fraction of a full fringe, because the base line is so small. It is therefore very difficult to defringe these data. You could try and use the tool "SWS_Defringe", which can be found under "Special Fns...". Please read the very detailed on-line help text for this tool. The FFT method won't work, because of the short baseline, but the sinus or DoK methods could improve the result. To check if you have residual fringes in your data you can read in the Relative Spectral Response Functions (RSRF, calibration files Cal25_xx, found in your ISAP installation under .../data/cal_sws), overplot those to the spectrum, and compare the frequency of the modulation. Note, that the unit of the cal files is microvolt/second/jy. So you will have to scale the RSRF to the level of your spectrum (which is in Jy). For more details please see the SWS Handbook.

The overplotting of the RSRF function can also help to address the question of "fake" features in the spectrum. See the handbook.

14) **Smooth**

This function reduces the noise in the flux values of an AAR (either in unique parts, such as detectors, or all together) by convolution with a selected function: Boxcar, Gaussian, Lorentzian, Median, Hanning, or Sinc. If you are interested in line profiles you should, however, be very careful when choosing the function width (in other words: don't smooth too much!). Stay well below the FWHM of your lines (or try using the boxcar or median filter).

15) **Loop over all AOT bands**

Store the processed spectrum and go back to step 5) to reduce the other lines in the same manner. Be aware of the plot style. Large data sets should be plotted in the "unconnected, without color preference" mode, which is the fastest (remember the "Ctrl-c" + "return" option, see step)

16) **Further steps**

After you have cycled through all lines in the way described above, the basic reduction of SWS02 data is complete. If you wish to combine all the single spectra to one complete spectrum again, you can do so: make one of the spectra the prime data set and then select the other bands one after the other ("Stored Data Sets") and use "Stored Data

Sets" →"Merge with Prime Data Set". This operation simply combines all selected data into one, larger AAR. A quicker way is provided by "Merge group of AARS" under "Special Fns...". Before you merge, you should consider to mask out in the overlap region of two lines (if they exist) the data of one of the lines, if these are of lower quality than the corresponding data in the other lines. Alternatively you could average across the overlap region (by using the line tag).

A further option of data processing is REBINning (see the web documentation). ARITHM may be helpful for all kinds of arithmetic operations (e.g. to subtract or add two spectra or to add a constant, etc.).

Finally you may wish to analyse the processed spectrum, e.g. measure some line fluxes. The "LINE FIT" tool, accessible via the tool box for selected data, provides all you need to fit a baseline and the line itself, it returns the fit parameters, line fluxes etc., and suggests several possible line identifications, based on the specified redshift (radial velocity) and a built in line list (including unidentified infrared bands – UIBs – and some unidentified lines).

The unit conversion tool (accessible under "Special Fns...") may also be of interest for you.

"Zodiacal" can be used to subtract the dominant foreground at SWS wavelengths. "Dered" (available in the command line mode only) allows to deredden the spectrum according to a given extinction law.

When you exit the GUI you will be in ISAP's command line mode. By typing *aar=gui_output()* you'll get the aar which was prime on exit of the GUI. You can then work on your aar just like in a normal IDL command line environment. You can go back to a GUI session any time by typing *start_gui* or *start_gui,aar_in=aar*. The latter command passes an aar from your command line session directly into the GUI (see the information on the ISAP web page under "IDL-ISAP interface").

If you need more help, you should contact (or consider to schedule a visit to) IPAC, MPE, VILSPA or one of the other national ISO data centers. Informations about IPAC, MPE or the VILSPA IDC can be found on the web:
http://www.ipac.caltech.edu/iso/
http://www.mpe.mpg.de/www_ir/ISO/
http://isowww.estec.esa.nl/
You can also email us any kind of problems with ISAP and any suggestions you have:
isap_dev@ipac.caltech.edu. (Within the GUI just click on "Send Mail")
The author of this recipe can be contacted via: sturm@mpe.mpg.de

If you are going to publish results which are based on a data reduction and/or analysis with ISAP we would very much appreciate the following acknowledgement:
"The ISAP Spectral Analysis Package (ISAP) is a joint development by the LWS and SWS Instrument Teams and Data Centres. Contributing institutes are CESR, IAS, IPAC, MPE, RAL and SRON."

And now: Good luck with the reduction and analysis of your ISO data!