ADAS602: Spectral line profile fitting

The program fits a theoretical spectrum composed of a polynomial background with several Gaussian shaped lines to an observed spectrum to obtain the positions, widths and fluxes of the observed lines. The maximum likelihood method is used to provide estimates of the line and background parameters, their standard errors and 95% confidence limits.

Background theory:

The representation is as a constant, linear or quadratic background together with several Gaussian shaped lines. The observed data are therefore fitted to the equation

$$I_k = b_0 + b_1 x_k + b_2 x_k^2 + \sum_{i=1}^{L} h_i \exp[-((x_k - xo_i)/w_i)^2] \qquad 7.2.1$$

where I_k is the observed count rate at a particular detector pixel position x_k. L is the number of ordinary lines in the spectrum to be fitted with xo_i the pixel position of the centre of the line i and w_i is its half width. Statistical variation in I_k is assumed to have the normal form. With variance proportional to the mean value of I_k. Hence

$$\sigma_k^2 = Var\, I_k = \sigma^2 E(I_k) \qquad k = 1,2,...,P \qquad 7.2.2$$

where σ^2 is an unknown constant of proportionality, P is the number of pixel positions and E denotes the expected value. Under these conditions, the likelihood function can be formed and values which maximise it also maximise its (natural0 logarithm. Thus to maximise the $\log$ likelihood to estimate the N variables $b_0, b_1, b_2, h_i\, xo_i, w_i$ the need is to minimise the function

$$\sum_{k=1}^{P} \left[\frac{I_k}{\sigma_k} - \frac{1}{\sigma_k}(b_0 + b_1 x_k + b_2 x_k^2 + \sum_{i=1}^{L} h_i \exp[-((x_k - xo_i)/w_i)^2] \right]^2$$

$$7.2.3$$

The σ_k^2 are unknown. Assuming however that the counts per sampling interval follow the Poisson distribution, the σ_k^2 are equal to the true unknown mean count per sampling interval divided by the number of sampling intervals. If the unknown mean count rate is estimated by I_k, then an estimate of σ_k^2 follows, namely I_k divided by the number of intervals where the number of intervals is the dwell time in seconds. A standard numerical routine will find the unconstrained minimum of the function 7.2.4, i.e. minimise

$$F(x) = \sum_{i=1}^{P} [f_i(x)]^2 \qquad 7.2.4$$

The functions $f_i(x)$ (the residuals), their first derivatives and a starting point must be supplied. The starting point was obtained from the observed spectrum by indicating the number of lines to be fitted and estimating their peak positions and full-width half-maxima. The background was also obtained from the spectrum and estimates of b_0, b_1, b_2 made. Once estimates of $b_0, b_1, ...$ (denoted by $\hat{b}_0, \hat{b}_1, ...$) have been found a second routine is used to find the diagonal elements of the variance-covariance matrix C. This routine depends on the use of the standard result of maximum likelihood theory which states that the variance-covariance matrix of the estimates is approximately given by the negative of the inverse of the expected value of the matrix of second derivatives of the logarithm of the likelihood. In the routine it is assumed that the Hessian (i.e. matrix of second partial derivatives) of $F(x)$ at the solution can be approximated adequately by $2J^T J$ where J is the Jacobian (i.e. matrix of first partial derivatives) of $F(x)$ at the solution. C is then given by

$$C = \hat{\sigma}^2 (J^T J)^{-1} \qquad\qquad J^T J \text{ on-singular} \qquad\qquad 7.2.5$$

where $\hat{\sigma}^2$ is the estimated variance of the residual at the solution ($\hat{\hat{x}} = \hat{b}_0, \hat{b}_1, \ldots$) and is given by

$$\hat{\sigma}^2 = F(\hat{x}) / (P - N) \qquad\qquad 7.2.6$$

The square root of the diagonal elements of C give the estimated standard errors of the corresponding elements of $\hat{x}$. A 95% confidence interval for each of the elements of $\hat{x}$ is obtained using $\hat{x}_j \pm t_{P-N;97.5} (C_{jj})^{1/2}$ where $t_{P-N;97.5}$ denotes the upper 97.5% point of the Student's t distribution with $P - N$ degrees of freedom. Note that the approximation of the Hessian by $2J^T J$ is justified as follows: the Hessian matrix $G(x)$ is of the form

$$G(x) = 2[J^T(x)J(x) + \sum_{i=1}^{P} f_i(x) G_i(x)] \qquad\qquad 7.2.7$$

where $G_i(x)$ is the Hessian matrix of $F(x)$. In the neighbourhood of the solution $\left\| f(x) \right\|$ is often small compared to $\left\| J^T(x)J(x) \right\|$, for example when $f_i(x)$ represents the goddness of fit of a non-linear model to observed data. In such cases $2J^T(x)J(x)$ may be an adequate approximation to $G(x)$. This avoids the need to compute or approximate second derivatives of the $f_i(x)$.

Program steps:

These are summarised in the figure below.

Figure 7.2
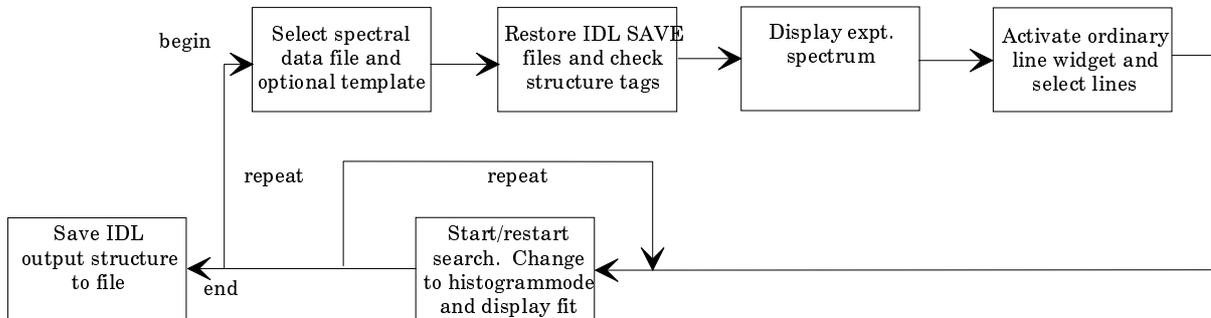
Interactive parameter comments:

Move to the directory in which you wish any ADAS created files to appear. These include the output text file produced after executing any ADAS program (*paper.txt* is the default) and the graphic file if saved (e.g. *graph.ps* if a postscript file)

The **file selection window** has the appearance shown below

1. The form of the primary input spectral data is selected at a). Input data sets are expected to be in the form of IDL 'SAVE' files. Forms allowed are as follow:

    a) The form *Standard* comprises a single array defined as

        spectrum :   *dblarr()*

        header :   *''*   or *structure (ie.={a:a, b:b,...})*

spectrum   contains the spectral counts as a function of pixel position. The array length must be <=1024. The array may be two-dimensional in which case the first index is the pixel number in the dispersion direction and the second the pixel number in the spatial (or other) direction. This latter case allows the cycling option described in section 5.2.

header   is an optional string of the user's choice or may be a structure

b)   The form *Processed* is that of the output structure from ADAS602 itself. This form allows re-entry of output data for re-analysis. The input file contains a structure *lstr* defined as:

lstr.b0 :	*0.0*	*lstr.eseb0:*	*0.0*
lstr.eclb0 :	*0.0*	*lstr.b1:*	*0.0*
lstr.eseb1 :	*0.0*	*lstr.eclb1:*	*0.0*
lstr.b2 :	*0.0*	*lstr.eseb2:*	*0.0*
lstr.eclb2 :	*0.0*	*lstr.w:*	*dblarr(10)*
lstr.esew1 :	*0.0*	*lstr.eclw1*	*0.0*
lstr.xo :	*dblarr(10)*	*esexo:*	*dblarr(10)*
lstr.eclxo :	*dblarr(10)*	*eclh:*	*dblarr(10)*
lstr.esewv :	*dblarr(10)*	*eclwv:*	*dblarr(10)*
lstr.flux :	*dblarr(10)*	*esef:*	*dblarr(10)*
lstr.eclf :	*dblarr(10)*	*getfit:*	*intarr(6)*
lstr.exptime :	*0.0*	*lstr.info:*	*0*
lstr.header :	*''*	*lstr.xmin:*	*0*
lstr.xmax :	*0*	*lstr.x:*	*indgen()*
lstr.y :	*dblarr()*	*lstr.ya:*	*dblarr()*
lstr.nsize :	*0.0*	*lstr.numline:*	*0*
lstr.nfxl :	*0*	*lstr.bchl:*	*0.0*
lstr.bchr :	*0.0*	*lstr.lineid:*	*strarr(10)*

with the meaning of the variables as follow:

b0   fitted constant part of the background

eseb0   estimated standard error in *b0*

eclb0   estimated 95% confidence limit in *b0*

b1   fitted linear part of the background

eseb1   estimated standard error in *b1*

eclb1   estimated 95% confidence limit in *b1*

b2   fitted quadratic part of the background

eseb2   estimated standard error in *b2*

eclb2   estimated 95% confidence limit in *b2*

w   fitted line widths

esewl   estimated standard error in *w* (fixed)

eclwl   estimated 95% confidence limit in *w* (fixed)

xo   fitted line centroid positions

esexo   estimated standard error in *x0*

eclxo   estimated 95% confidence limit in *x0*

h   fitted line heights at centroid

eseh   estimated standard error in *h*

eclh   estimated 95% confidence limit in *h*

|	
esewv	estimated standard error in *w* (variable)
eclwv	estimated 95% confidence limit in *w* (variable)
flux	fitted line flux (counts under profile)
esef	estimated standard error in *flux*
eclf	estimated 95% confidence limit in *flux*
getfit	options selected (see code)
exptime	the exposure time (1.0 if expval not set)
info	error from MINPACK routine (see XXDER1)
header	comment
xmin	minimum pixel for fitting range
xmax	maximum pixel for fitting range
x	pixel values
y	spectral data
ya	fit to data
nsize	spectral range width
numline	number of fitted lines
nfxl	number of fitted lines with fixed positions
bchl	initial estimate of background
bchr	inistial estimate of background
lineid	preliminary line identification labels

c) The form *Other* is not yet implemented and we are open to suggestions for it (eg. and ascii tabulated version of *lstr* os *spectrum/header*?)

2. Data root b) shows the full pathway to the appropriate data subdirectories. Click the *Central Data* button to insert the default central ADAS pathway to */../adas/adas/arch602*. Note that there are only some sample data in this subdirectory (viz. a standard data set '*...clean*' and a processed data set '*fit....*') as there is no formal central ADAS data of archive type. Click the *User Data* button to insert the pathway to your own archived data for ADAS602.

3. The Data root can be edited directly. Click the *Edit Path Name* button first to permit editing. Available sub-directories are shown in the large file display window b). Scroll bars appear if the number of entries exceed the file display window size.

4. Click on a name to select it. The selected name appears in the smaller selection window c) above the file display window. Then its sub-directories in turn are displayed in the file display window. Ultimately the individual data-files are presented for selection.

5. Clicking on the *Browse Comments* button displays the information header associated with the selected data-file.

6. A second data set may be selected at d), e) and f). Such a data set must be in the form of a processed output data from ADAS602. Its purpose is to provide initial selection values for analysing the new data set. This facility is important for repetitive analysis of data from the same spectral region where the same set of lines is being picked out. Note that the default pathway is to the */.../adas/pass* directory in this case. Note also that if a processed file is selected **it will be used** to fit the new data.

7. Once a primary data file is selected, the buttons at the bottom of the main window become active.

8.   Clicking the *Done* button moves you forward to the next window.   Clicking the *Cancel* button takes you back to the previous window

```
┌──────────────────────────────────────────────────────┐
│ ▽            ADAS602 INPUT                              │
│  ┌──────────────────────────────────────────────────┐  │
│  │ Select Data Form:    ◇  Standard  ◇ Processed ◇ Other │ │
│  └──────────────────────────────────────────────────┘  │
│  ┌──────────────────────────────────────────────────┐  │
│  │              Input Dataset                    a)   │  │
│  │  Data root   [/disk2/brooks/cds/may_97/avg5/  ▼]   │  │
│  │                                                b)  │  │
│  │  [Central data]   [User data]   □ Edit Path Name   │  │
│  │                                                    │  │
│  │        [ s6882r00_n2_clean5              ]         │  │
│  │        ┌──────────────────────────────┐ △         │  │
│  │        │  . . .                        │           │  │
│  │  Data File: s6882r00_n2_clean5        ▼│           │  │
│  │        │  . . .                        │   c)      │  │
│  │        └──────────────────────────────┘ ▽         │  │
│  │  [Browse Comments]                                 │  │
│  └──────────────────────────────────────────────────┘  │
│  ┌──────────────────────────────────────────────────┐  │
│  │           Select PROCESSED Template                │  │
│  │  Data root   [/disk2/brooks/adas/pff/adas602/ ▼]   │  │
│  │                                                d)  │  │
│  │  [Central data]   [User data]   □ Edit Path Name   │  │
│  │                                                    │  │
│  │        [                              ▼ ]          │  │
│  │        ┌──────────────────────────────┐ △  f)     │  │
│  │        │                               │           │  │
│  │  Data File: │                     ▼    │           │  │
│  │        │                               │           │  │
│  │        └──────────────────────────────┘ ▽         │  │
│  │  [Browse Comments]                          e)     │  │
│  └──────────────────────────────────────────────────┘  │
│  [Cancel]   [Done]                                      │
└──────────────────────────────────────────────────────┘
```

The **processing options window** has the appearance shown below

9.   The spectrum (counts Vs pixel number) is displayed at a) and a title for the hard copy of the graph may be entered at b).

10.   In general, one wishes to select only a segment of the spectrum for fitting and to adjust the y-scale. This is done by entering appropriate values at c). Note that the entries have effect immediately so the graph can be adjusted rapidly to requirements. Line fitting choices are made at the right side of the screen. At d) select the number of line features you wish to fit to the spectrum. A maximum of ten is allowed.   At e), select the exposure time if the data are in counts. Omit this if they are in counts/sec.

11.   It is possible to constrain the position of any of the line features in the fitting procedure. To do so, click the *Fix Line Positions* button at f) and enter the pixel positions at which you wish the features locked. It need not be all of the features.

12.   A facility is also present to attach a name to a line feature which is helpful for known lines. These names are carried through to the output data structure. Click the *Identify Lines* button at g) and enter the names in the appropriate fields below.

13.   Control of line widths is provided at h). Select first of all whether you wish the line widths to be permitted to vary from the initial values during the fitting or not. If not is selected, a common line width must be entered.   Then chose if you wish all the lines to have the same (but varying) width or if each line width can vary separately.

14. At i) select the form of background shape you wish to allow. Note that if you prepare the data with no background then the fitting to no background is allowed by deactivating the *Background Options* button

15. The way the experimental pixel count data is weighted during the fit optimisation may be chosen at j). The *Weighted Fit* button weights a count with the square root of the count. The *Equal Weighted Fit* button imposes the same weight on the data at each pixel position.

16.
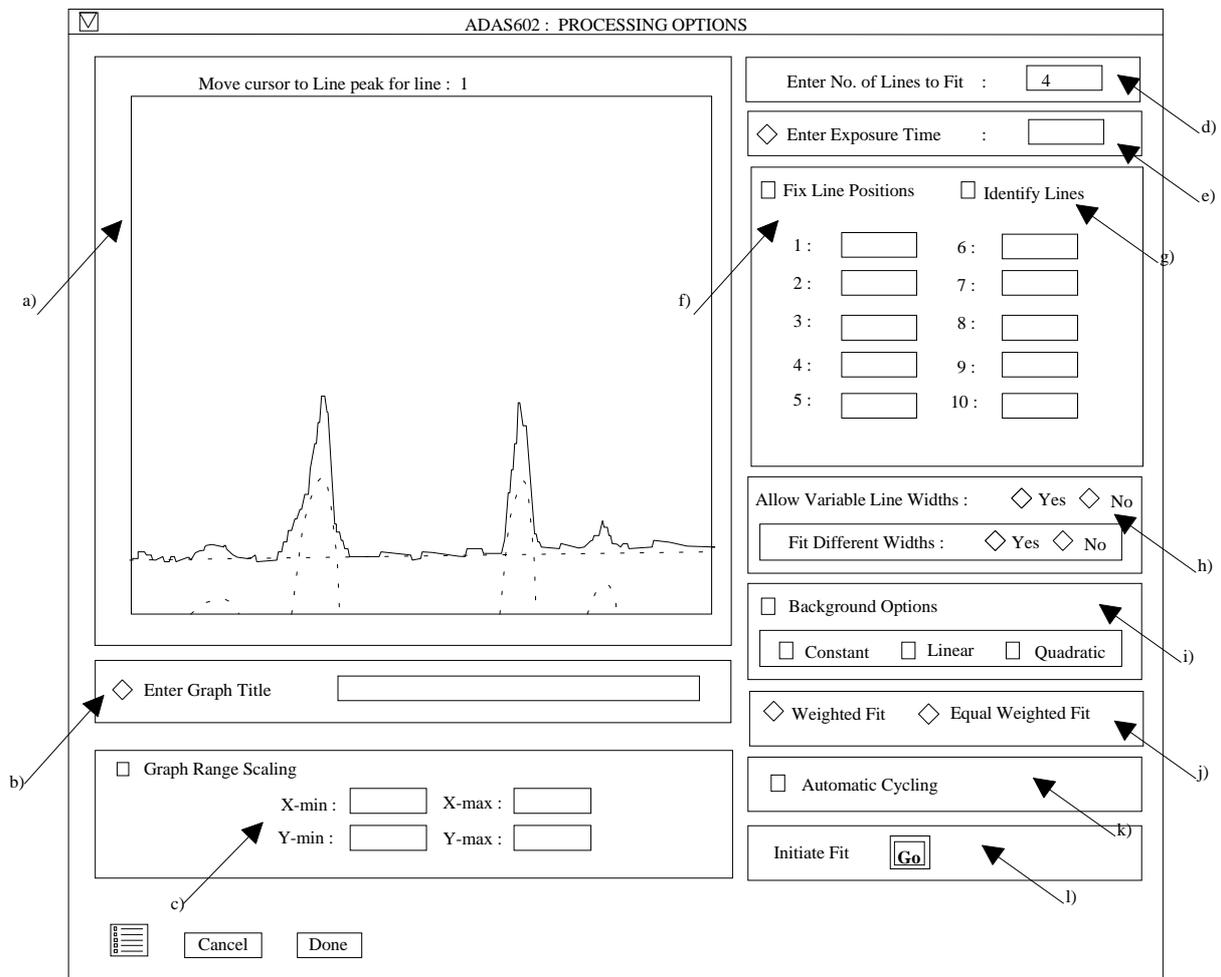
17. If there are multiple spectra in the input data set, then the fitting procedure can be made to cycle automatically through each data set in turn using the initial settings from the first fit. Activate the *Automatic Cycling* button at k)

18. Finally click the button '*Go*' at l) to initiate the fitting process. The displayed graph changes at this point to histogram mode. The mouse cursor position on the graph screen becomes meaningful and the left and right mouse buttons active. Advisory information of the action currently expected is given above the graph depending on the processing options chosen. In general one clicks to mark the peak, left half width position and right half width position for each feature in turn. A small square marks each selected position. Finally one marks the lowest and highest levels of the background. Clicking the right button can restart the process from the beginning. After completing the required marking, the fitting optimisation takes place. The figure is redisplayed with the fitted Gaussians and background superimposed in dotted outline.

19. At the base of the window, the icon for *Exit to Menu* is present. This quits the program and returns you to the ADAS series 6 menu. Remember that *Done* takes you forward to the next screen while *Cancel* takes you back to the previous screen.

The **output options window** appearance is shown below

20.   As in the previous window, the full pathway to the file being analysed is shown for information. Also the *Browse comments* button is available.

21.   A graph hard copy is available by activation of the *Graphical Output* button a). The File name box then becomes editable. If the output graphic file already exists and the *Replace* button has not been activated, a 'pop-up' window issues a warning.

22.   A choice of output graph plotting devices is given in the Device list window at d). Clicking on the required device selects it. It appears in the selection window above the Device list window.

23.   The *Text Output* button activates writing to a text output file. The file name may be entered in the editable File name box when *Text Output* is on. The default file name 'paper.txt' may be set by pressing the button *Default file name*. A 'pop-up' window issues a warning if the file already exists and the *Replace* button has not been activated.

24.   At e) output of the processing and analysis may be saved by clicking the *Passing File* button. The *Default file name* sends the output file to your */../adas/pass* directory as *adas602.pass*. Note that the output file is an IDL SAVE file structured in the manner described above. Note that defaults are updated after each run.

25.   At the base of the window, the icon for *Exit to Menu* is present. This quits the program and returns you to the ADAS series 6 menu. Remember that *Done* takes you forward to the next screen while *Cancel* takes you back to the previous screen.

Illustration:

The example shows the fitting of a section of a cleaned spectrum from the CDS-NIS spectrometer on the SOHO-satellite. The table shows the statistical analysis of the fit parameters found.

Figure 7.02a

ARBITRARY COUNTS VS. PIXEL NUMBER
ADAS   : ADAS RELEASE: ADAS98 V2.1 PROGRAM: ADAS602 V1.5 DATE: 26.03.99 TIME: 08:31
FILE   : /packages/adas/adas/arch602/s6882r00_n2_clean5 : STANDARD DATA
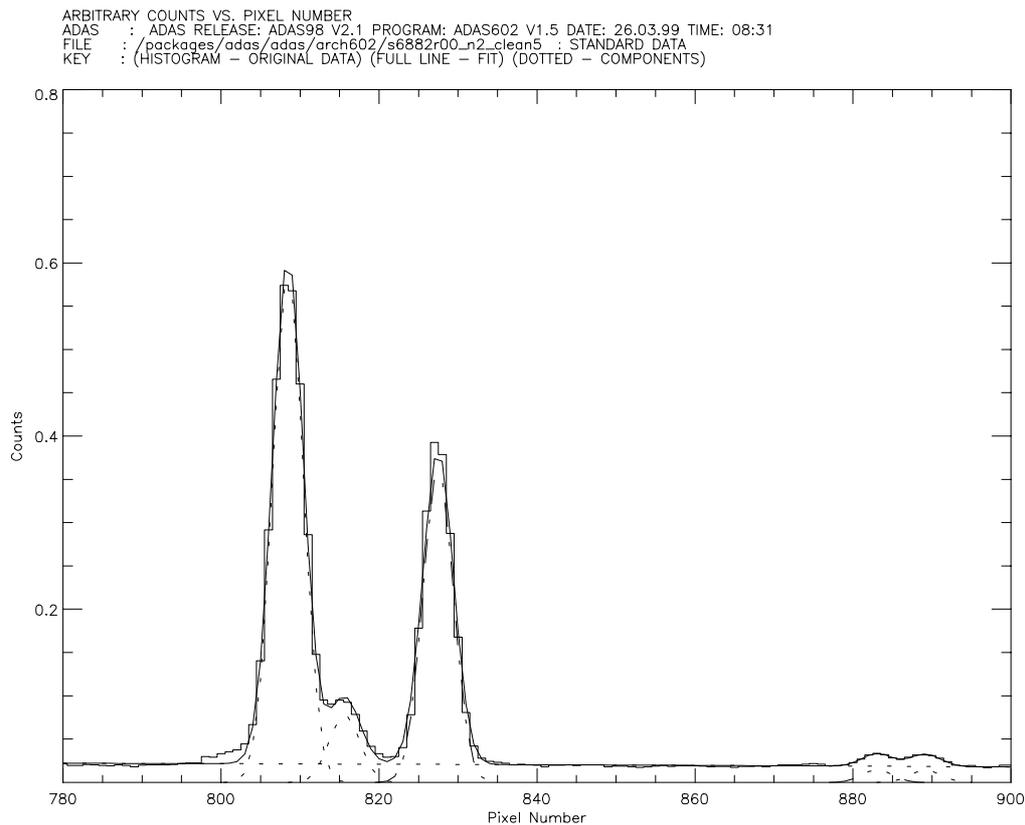KEY   : (HISTOGRAM – ORIGINAL DATA) (FULL LINE – FIT) (DOTTED – COMPONENTS)

Table 7.02a

```
     ADAS RELEASE: ADAS98 V2.1 PROGRAM: ADAS602 V1.5 DATE: 26.03.99 TIME: 08:31
*** TABULAR OUTPUT FROM MAXIMUM LIKELIHOOD SPECTRAL LINE PROFILE FITTING PROGRAM: ADAS602 - DATE: 26.03.99
***

 DATA FILENAME : /packages/adas/adas/arch602/s6882r00_n2_clean5 DATA TYPE : STANDARD
      TEMPLATE : NONE

--------- LINE WIDTH OPTIONS ---------
-------------------------------------
 ALL LINE WIDTHS FIXED ON INPUT   : N
 VARIABLE LINE WIDTHS WERE FITTED : N

--------- BACKGROUND OPTIONS ---------
-------------------------------------
 CONSTANT COMPONENT  : Y
 LINEAR COMPONENT    : Y
 QUADRATIC COMPONENT : Y

------------- OTHER OPTIONS --------------
 WEIGHTED FIT        : Y
 INFO ERROR FLAG FROM MINPACK ROUTINE : 1
 EXPOSURE TIME : 1.00000
 NO. OF LINES  : 5              NO. OF LINES WITH FIXED POSITIONS : 0


 KEY
 -----
 ESE - ESTIMATED STANDARD ERROR (%)
 ECL - ESTIMATED 95% CONFIDENCE LIMIT (%)
--------------------------------------------- OUTPUT TABLE ---------------------------------------------
-

 DATASET : 1
 --------------
 BACKGROUND
 ------------
 CONSTANT  : -0.0368380       ESE : 519.882        ECL : 1031.62
 LINEAR    : 0.000171724      ESE : 267.236        ECL : 530.286
 QUADRATIC : -1.23206e-07     ESE : 222.799        ECL : 442.108
 INDEX   POSITION                      HEIGHT                              FLUX
WIDTH                    IDENTIFICATION
        (XO)      (ESEXO)    (ECLXO)    (H)      (ESEH)    (ECLH)               (ESEF)    (ECLF)
(ESEW)    (ECLW)
------------------------------------------------------------------------------------------------------
------------------------------------------------
   1    808.461    0.003      0.007      0.586    1.441     2.860     2.854    1.691     3.355
2.747      0.884      1.755         u
```

```
   2    815.624    0.014    0.027    0.078    4.403    8.738    0.378    4.491    8.912
2.747      0.884    1.755       u
   3    827.467    0.004    0.008    0.363    1.790    3.552    1.770    1.996    3.961
2.747      0.884    1.755       u
   4    883.053    0.042    0.083    0.015   14.571   28.914    0.071   14.598   28.967
2.747      0.884    1.755       u
   5    889.164    0.043    0.085    0.014   15.127   30.016    0.070   15.152   30.068
2.747      0.884    1.755       u
------------------------------------------------------------------------------------------------
------------------------------------------------
```

Notes: