Chapter 5

Analysis of Observational Data from the SOHO Spacecraft

5.1 Introduction and Objectives of the Analysis

Our studies focus on the temperature/density structure of the solar atmosphere and its dynamic nature, taking advantage of the high spatial, temporal and spectral resolution of the CDS and SUMER spectrometers. The first stage is to use the most accurate derived atomic data, produced by ADAS, and couple it with calibrated line intensity measurements to investigate the differential emission measure distribution from the chromosphere through to the corona. The high quality of the atomic data will allow us to investigate and quantify the sources of error associated with this type of analysis more precisley than here-to-fore and determine whether it is a realistic approach for studying the structure of the solar atmosphere. The solar atmosphere's clearly dynamic nature leads us to believe that the technique may be limited. Subsequently, we wish to synthesize dynamic solar models linked to the observations by making geometrically valid spectrometer line of sight calculations to use in predictive studies.

In a preliminary study of the observations we develop critically the analysis techniques which will be used in all future work. Thus we seek to gain an understanding of the analysis problems and the peculiarities of the instruments. Also, it is essential to do this to provide feedback for refinement of the observing sequences. In this chapter, we develop techniques of data handling and analysis that will be used principally for computing the differential emission measure with line intensities obtained by both CDS and SUMER. In so doing, we test key software components for the DEM analysis and propose baseline software modifications to deal more effectively with various aspects of the observational data. A critical study is also made of the statistical variability of the spectrum line intensities obtained by the CDS spectrometer. This latter study follows that applied to the CHASE spectrometer which flew aboard Spacelab 2 in 1985 (Lang, Mason & McWhirter (1990)). Finally, a reconstruction of the DEM using CDS-NIS data is presented to demonstrate the validity of the analysis procedures and the coupling of the different aspects of the problem. In doing this, some comment is made on the sensitivity of the normal incidence detectors. The presentation is however preliminary. The intensity calibration issue for CDS-NIS remains undecided and so our definitive differential emission measure study remains on 'hold'. However, reexamination of potential blended lines, back checks on the atomic physics of particular lines etc. have been commenced on the basis of this preliminary work.

5.2 Spectral Line Profile Fitting Procedure

Central to the aims of our work is the analysis of spectral line intensities. To obtain the fluxes observed in the lines we fitted a theoretical spectrum to the observed one. The method used is to fit multi-Gaussian line profiles and a background using a maximum likelihood method together with statistical analysis to calculate the total integrated counts under the line profiles along with their estimated standard errors and 95% confidence limits. The basic procedure is due to Lang et al.(1990) and was used in their analysis of data from the CHASE experiment. It has been adapted here to extend its capabilities while retaining its simplicity and flexibility. A major modification was to integrate the routine to both the SOHO-CDS and SOHO-SUMER data acquisition software. In this section a brief outline of the statistical methods used is presented for completeness followed by an overview of the modifications, associated tests and capabilities of the latest IDL version.

5.2.1 Overview of Line Fitting Method

A full explanation of the statistical methods used was presented by Lang et al.(1990). Here we restrict ourselves to outlining the main features and note the different observational mechanisms of CDS compared with CHASE. CHASE built up a profile of a particular line at a number of grating positions while CDS gives the whole spectrum at one position. The observed spectrum is represented by the form below,

$$I_k = b0 + b1x_k + b2x_k^2 + \sum_{i=1}^{L} h_i \exp - \left(\frac{(x_k - xo_i)}{w}\right)^2 \tag{5.1}$$

where I_k is the observed counts at integer bins, L is the number of lines to be fitted, b0 is the flat component of the background, b1 introduces a linear slope to that background, b2 introduces a curve to that background, x_k gives the pixel positions, h_i gives the peak of each line, w is the width to fit for each line and xo_i is the fitted positions of the lines which may be to fractions of a pixel. The statistical variation in the observed counts in the line, I_k, is assumed to have a normal distribution with variance proportional to the mean. The true mean at any I_k is unknown but if it is equal to the expected value of I_k, denoted $E(I_k)$, then an estimate of the variance of I_k follows from $\sigma_k^2 = \sigma^2 E(I_k)$. Here, σ^2 is an unknown constant of proportionality and σ_k^2 is the unknown statistical variance. As defined by Lang et al(1990) the natural logarithm of the likelihood function is

$$\log L = \sum_{k=1}^{P} - \log \sigma_k (2\pi)^{1/2}$$

$$-\frac{1}{2}\sum_{k=1}^{P}\frac{1}{\sigma_k^2}\left[I_k - \left(b0 + b1x_k + b2x_k^2 + \sum_{i=1}^{L} h_i \exp - \left(\frac{(x_k - xo_i)}{w}\right)^2\right)\right]^2 \tag{5.2}$$

where k=1,...,P and P is the number of pixels in the range of the data. We maximise the likelihood function by maximising eq.5.2. That is, minimise the negative part of the right hand side of eq.5.2,

$$\sum_{k=1}^{P}\left[\frac{I_k}{\sigma_k} - \frac{1}{\sigma_k}\left(b0 + b1x_k + b2x_k^2 + \sum_{i=1}^{L} h_i \exp - \left(\frac{(x_k - xo_i)}{w}\right)^2\right)\right]^2 \tag{5.3}$$

In Lang et al.(1990) this was done by a NAG library routine. However, we removed the dependence on NAG for portability and compatibility with ADAS by using an adapted version of the MINPACK routine XXDER1. This minimises a sum of squares of functions provided, so we write eq.5.2 in the form $F(x) = \sum_{i=1}^{P} [f_i(x)]^2$ and specify the functions $f_i(x)$ along with their first derivatives. An initial estimate must be supplied and the code then iterates to find the solution that minimises eq.5.3. The parameters required for this initial estimate are dependent on options outlined below but usually include the number of lines, their centroid positions, estimated widths and an expected background level. The code obtains estimates of the unknown quantities, σ_k, by dividing the estimated value of I_k by the exposure time. This was retained for compatibility with CHASE. Finally, the estimated standard error of the individual quantities (e.g. b0, b1, b2, w etc.) and estimated 95% confidence limits are computed by standard methods of maximum likelihood. See Lang et al.(1990) and Lang & Payne (1987) for further description.

5.2.2 Integration with Existing Instrument Software

In the subsequent text we refer to the line profile fitting routine by its name, *mlcds*. To merge *mlcds* with the CDS and SUMER software it was necessary to develop first an IDL interface. This is because all the CDS software is written in IDL and the data, although archived in FITS format, is accessible only through this path. In addition, there are plans to archive the SUMER data in the same format and using the same organisation as that for CDS. This has not yet been done. The present recommended method for obtaining corrected SUMER data is directly from the CD ROM disks relayed to the Max Planck Institute in Lindau, Germany. The form of these data are IDL save/restore files and so are easily manageable by the IDL interface as for CDS. With this last point in mind, the original strategy was to take the SUMER data and streamline it into the form that *mlcds* expected from the way it was used for CHASE. Thus we had to specify an array containing the spectrum, the minimum and maximum pixel numbers over the range of the spectral segment and the exposure time of the observations. The exposure time was needed for CHASE as the count rates were recorded in quarter seconds and so had to be readjusted. Here, it is just used to divide the data to produce quarter seconds, before passing to the original code, and is then used to multiply up again at the end. For DEM studies the exposure time is redundant as the data are calibrated and come as a count rate. Therefore, the option of specifying the exposure time was made into a keyword (/exptime) to allow flexibility between the two cases.

We sought to retain full use of the available baseline CDS and SUMER software. Therefore, we returned all the output information from the fitting program in the form of a structure with individual tags referencing the information. The original data and references to the source were maintained within this organisation to allow immediate reconstruction of the spectrum. The contents of the structure and their meaning are given in table.5.1.

5.2.3 Illustrative Examples and Outline of Analysis Options

During the course of the investigations presented in the next sections it became necessary to make a number of improvements to the original program to extend its' applicability. Table.5.2 provides an overview of the IDL keywords currently implemented. The most important capabilities, are as follows. *mlcds* can use an initial fit to one line to fit automatically the same line repeatedly through multiple datasets. For the extremely large datasets generated by the observing sequences it is unrealistic to treat them all individually, especially if the data are of regions where the behaviour of the emission is expected to be similar. A switch is provided (/autoloop) which invokes the repeat facility. The ability to fit individually is of course retained and is most useful for lines that show dynamic behaviour, for example. The (/oldstr) keyword allows specification of an initial structure, normally taken from the fit to another dataset, that can then be used to supply the information for the current fit, to the code. This was implemented, during a visit to MPAë Lindau, where it is usual to do automatic fitting as a background job. /oldstr works in conjunction with the /noprompt keyword for background jobs.

A number of routes in the analysis are allowed following user response to specific prompts from the code. After start up, the number of lines for fitting is entered, to a maximum of ten, and then the number with fixed centroid positions is specified.

Variable	Type	Meaning
B0	Floating	Flat part of background
ESEB0	Floating	Estimated standard error in B0
ECLB0	Floating	Estimated 95% confidence in B0
B1	Floating	Sloped part of background
ESEB1	Floating	Estimated standard error in B1
ECLB1	Floating	Estimated 95% confidence in B1
B2	Floating	Curved part of background
ESEB2	Floating	Estimated standard error in B2
ECLB2	Floating	Estimated 95% confidence in B2
W	Double Array(10)	Fitted widths of lines
ESEWL	Floating	Estimated standard error in W(fixed)
ECLWL	Floating	Estimated 95% confidence in W(fixed)
XO	Double Array(10)	Fitted positions of lines
ESEXO	Double Array(10)	Estimated standard error in XO
ECLXO	Double Array(10)	Estimated 95% confidence in XO
H	Double Array(10)	Fitted heights at centroid of lines
ESEH	Double Array(10)	Estimated standard error in H
ECLH	Double Array(10)	Estimated 95% confidence in H
ESEWV	Double Array(10)	Estimated standard error in W(variable)
ECLWV	Double Array(10)	Estimated 95% confidence in W(variable)
FLUX	Double Array(10)	Counts under line profile
ESEF	Double Array(10)	Estimated standard error in flux
ECLF	Double Array(10)	Estimated 95% confidence in flux
GETFIT	Integer Array(6)	Options selected(see next table)
EXPTIME	Floating	Exposure times
INFO	Integer	Error flag from MINPACK routine
HEADER	String	Editable comment
XMIN	Integer	Minimum pixel of spectrum
XMAX	Integer	Maximum pixel of spectrum
X	Integer Array(nsize)	Pixels
Y	Double Array(nsize)	Spectrum
YA	Double Array(nsize)	Fitted spectrum
NSIZE	Long	Range of pixels
NUMLINE	Integer	No. of lines fitted
NFXL	Integer	No. lines with fixed positions
BCHL	Floating	Initial guess at background
BCHR	Floating	Initial guess at background
LINEID	String Array(10)	Preliminary line identifications

Table 5.1: Contents of the ‘Fitted’ structures returned from *mlcds*.

Keyword	Use	Default
EXPTIME	Specify the exposure time for the data and divide by it.	Off.
DATA	Specifies the spectrum array for one dataset.	Either data or loopdata must be set.
XMIN	Specifies the minimum pixel of the data range.	Left hand range of plot window.
XMAX	Specifies the maximum pixel of the data range.	Right hand range of plot window.
PSYM	Allows alteration of the plot type.	Histogram.
AUTOLOOP	Allows automatic processing of fits to many datasets.	Off.
LOOPDATA	Specifies the spectrum array for many datasets.	Either data or loopdata must be set.
HEADER	Allows entry of a comment (usually source data and username).	Empty.
NOBACK	No background in the data.	Background present.
NOPROMPT	Omit *Print to screen* prompt.	Prompt given.
POST	Make a postscript file of plot.	No postscript file.
OLDSTR	Fit data using a previous set of parameters.	New fit.

Table 5.2: Keyword options implemented in mlcds.

jl-ju	M+G (A)	Pixel^V	Flux^V	ESE^V	Pixel^F	Flux^F	ESE^F
1-2	1174.933	836.26	27440.1	2.2	836.26	27382.0	2.7
0-1	1175.263	844.14	22061.2	2.6	844.09	22618.3	3.0
1-1	1175.590	850.99	15648.1	11.3	851.85	19560.7	5.8
2-2	1175.711	854.70	73526.8	2.5	854.72	68494.8	2.2
1-0	1175.988	861.40	22520.5	2.7	861.30	22761.1	3.3
2-1	1176.370	870.36	25240.6	2.4	870.36	25372.7	3.0

Table 5.3: The results of the test on the SUMER CIII multiplet data of the alteration to allow fixed line positioning. M+G denotes Moore(1993 - Edited by Gallacher) of the wavelengths used. The superscript V refers to the variable positions and F to the fixed ones. The estimated standard errors are given as percentages of the total fluxes.

This latter option is essential to allow use of laboratory wavelengths in defining the separations of overlapped components of multiplets, for example. Before entering the fixed positions of the lines, an option is available to make a preliminary identification of them. This is recorded in the output data structure for use in subsequent analysis routines. It is envisaged that the fixed position option will be used mainly with known wavelengths but a facility to correct wrongly entered values or to confirm estimates of line centres is available at this point. The introduction of lines with fixed positions required a separation of the final term in eq.5.1 as,

$$\sum_{i=1}^{L-nfxl} h_i \exp - \left(\frac{(x_k - xo_i)}{w} \right)^2 + \sum_{i=L-nfxl+1}^{L} h_i \exp - \left(\frac{(x_k - xo_i)}{w} \right)^2 \tag{5.4}$$

where the variables are defined as before excepting *nfxl* which is the number of fixed lines. This modification was coded into the original FORTRAN.

Extensive tests were carried out on a series of selected multiplets comparing the derived intensity values with the expected values from ratios of transition probabilities and making a comparison of the variation in the errors of the fits. As an example, we used the data from the SUMER study OPAC_1. This sequence targets specific multiplets and tracks their variation in intensity as the instrument rasters across the solar limb. The component line intensity ratios arising from transitions from the same upper level alter due to the effects of opacity. We used the CIII($2s2p\ ^3P - 2p^2\ ^3P$)

transition in a dataset recorded on the solar disc. The dataset was chosen by viewing movies of limb brightening to confirm the selected one was indeed on the disc and to ensure the data was free from optical distortion. We first fitted the six members of the multiplet by allowing the line positions to vary, all lines to be the same width and the background to have a linear slope. This gave pixel positions for the six lines. From the graph of pixel versus wavelength a straight calibration line was drawn. Then best fixed pixel positions of the lines were estimated. The total counts of the lines found at the fixed positions were then obtained using the fitting program. The transitions and wavelengths, fitted pixel positions, fitted fluxes and estimated standard errors are given in table.5.3. The fit parameters are given for the case of fixed and variable line positions. The fixed positioning reduced the uncertainty for the blended pair 1-1 and 2-2 (1175.590$\mathring{A}$ and 1175.711$\mathring{A}$ or 851.85 and 854.72 fixed pixel position respectively) from 11.3% and 2.5% to 5.8% and 2.2% respectively, at the expense of about 0.5% accuracy in the others. An example plot of the fixed position fit to this multiplet is given in fig.5.1. We then compared the theoretical optically thin branching ratios with the observed intensity ratios obtained by the fitting program. The theoretical A-values were taken from ADAS which contains the best currently available data (Allard et al.(1990)). We estimate that these A-values are uncertain to at least 5%. Then, the theoretical ratio $I(2-2)/I(1-2)$ should be 3.00 ±0.21 and the ratios $I(0-1)/I(1-1)$, $I(2-1)/I(1-1)$ and $I(2-1)/I(0-1)$ should be 1.34±0.09, 1.67±0.12 and 1.250.09 respectively in optically thin plasma. The results for free positions give 2.68±0.09 for $I(2-2)/I(1-2)$, 1.41±0.16 for $I(0-1)/I(1-1)$, 1.61±0.19 for $I(2-1)/I(1-1)$ and 1.14±0.04 for $I(2-1)/I(0-1)$. While the results of the fixed positions give 2.50±0.09 for the $I(2-2)/I(1-2)$, 1.16±0.08 for $I(0-1)/I(1-1)$, 1.30±0.08 for $I(2-1)/I(1-1)$ and 1.12±0.05 for $I(2-1)/I(0-1)$.

The results of this test on the fitting program of fitting lines at fixed positions was satisfactory. The fixed position option was more accurate in evaluating fluxes from the blended lines as expected. In comparison with the theoretical line intensity ratios the results with the free fit were slightly better but less accurate. Intensity ratios calculated from the fixed position fits to the other lines had errors comparable with the free fit and were about the same at matching theoretical line intensity ratios.

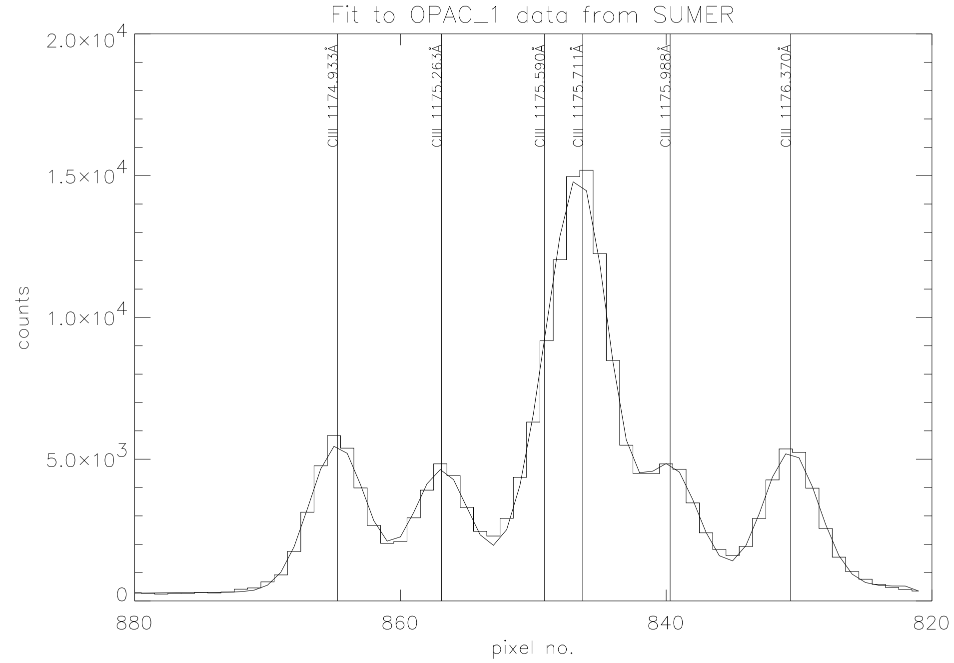


Figure 5.1: Example of fixed position fit to the CIII multiplet in OPAC_1 SUMER data.

Such success with this multiplet is essential for full analysis of opacity with the OPAC data.

During the other tests on CDS data the main problem arising was that blended lines were often closer together than the spectral resolution of the spectrometer i.e. they were often within one data pixel while the data were binned at one pixel. This is a fundamental limitation in the data that affects our ability to separate the lines. However, for optically thin plasmas, calculation of theoretical line ratios using ADAS together with spectral fitting of well separated lines will allow progress to be made in identification and inference of the make up of overlapped line intensities.

mlcds requires an estimate of the widths of each line. This is done using the mouse cursor by clicking on the left and right hand sides of the lines at about full width half maximum. In fact, the calculated widths are not the full widths at half maximum but are $\sim \frac{3}{5}$ of them (Lang (1996) - private communication). The treatment of the line widths has been made quite flexible for application to differing data requirements. An option is available to fix the widths of every line to be the same (answer 0(=No) to *line widths to vary ?*). Alternatively, one can allow the program to fit the widths but make them all the same (answer 1(=Yes) to *line widths to vary ?* and 0(=No) to *Fit different widths to lines*) or to fit them, but allow them to be different (line widths to vary and fit different widths to lines). These changes required alterations to the FORTRAN code to re-order the vectors containing the free parameters but the final terms in eq.5.1 are still those of eq.5.4. The vector w will contain the values of the variable widths or the widths that the program finds, but when using fixed widths for all lines it becomes a scalar quantity. In this case the exponential terms are constant and their derivatives drop out.

The ability to fit variable line widths was introduced because of a preliminary analysis on explosive event data from SUMER. The explosive components initially appear weak and then grow in intensity to dominate the central line. As first pointed out by Innes et al.(1997) these events are probably bi-directional jets of plasma created following magnetic reconnection in the chromosphere. Fig.5.2 shows the line width variation with time, along with intensity and position, in observations taken by SUMER in May. The solid line is the central unshifted non-explosive component

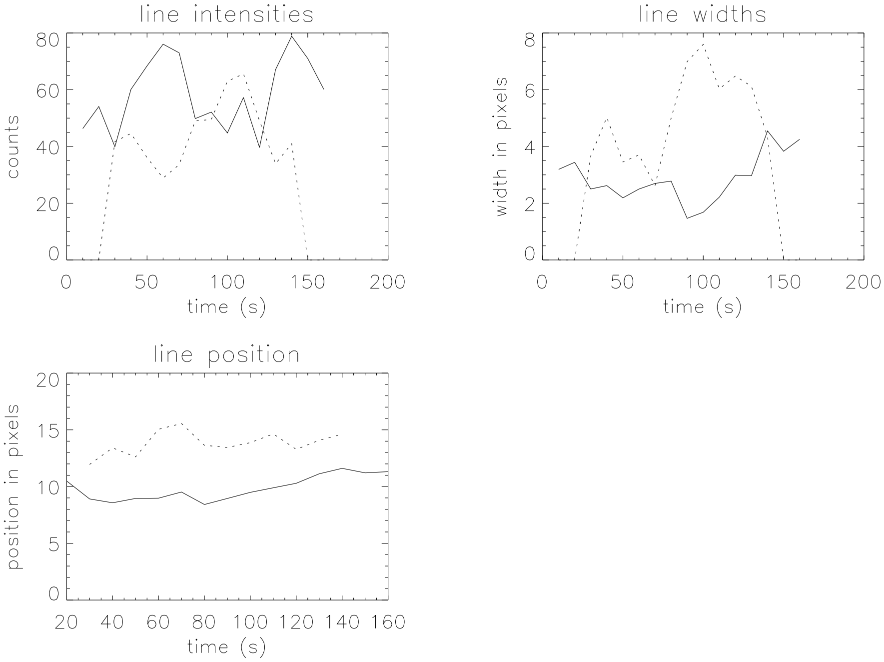


Figure 5.2: Example of application of variable widths option to SUMER explosive event data. In each case the solid curve refers to the stationary component and the dotted curve refers to the explosive component.

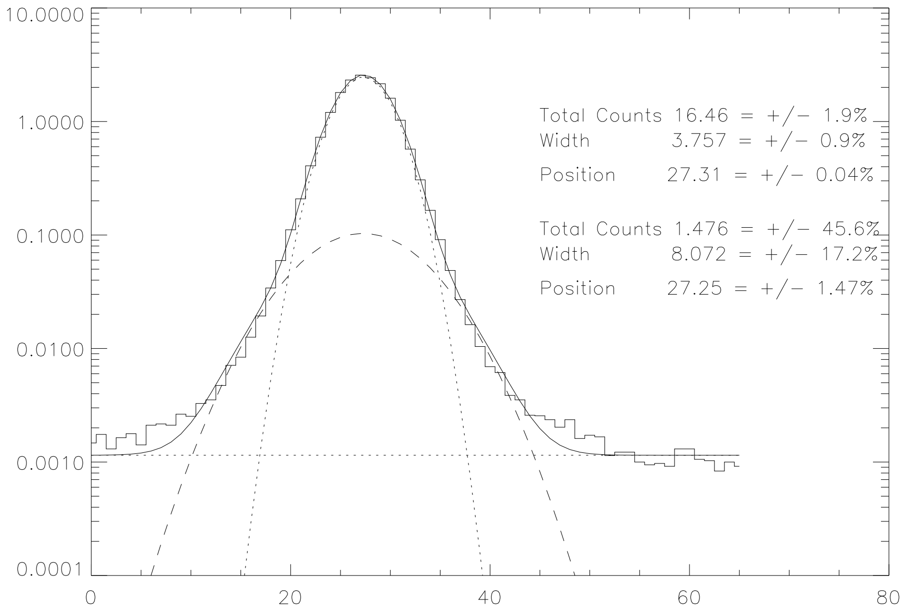


Figure 5.3: Example of variable width fit of two Gaussians to an OVI line taken while scanning above the limb. The x-coordinate is the pixel number in the wavelength direction. The y-coordinate is the number of counts.

and the dotted line is the explosive component. Note that the central component width remains relatively constant, between ∼3-4 pixels, as does the central position. The non-explosive component intensity appears to fluctuate substantially as the explosive component develops. This is due principally to imperfect separation of the fitting program at each time. The width of the explosive component varies over a much larger scale, ∼0-8 pixels, than the stationary component. This example illustrates the effectiveness of the variable width option in separating two components of different width. Another example is shown in fig.5.3. SUMER data from a scan off the limb is used to separate double Gaussians thus giving the contribution to the line from thermal broadening. The line used was OVI 1032$\AA$. Once again the variable width option was effective. Despite the much smaller number of counts the fit is within 45%.

Three options are available for specifying the shape of the background for which *mlcds* needs two initial estimates. These estimates are used to calculate the linear slope or curve to be added to the flat background dependent on the background choice. Once again the vector containing the free parameters had to be reordered and the various contributions to the $b0 + b1x_k + b2x_k^2$ part of eq.5.1 are dropped out dependent on the user choices.

The fitting procedure normally performs a fit to the counts in each line using a weight function which is the square root of the number of counts. In the case of the GIS 'fixed' patterns we found that this was not satisfactory. We implemented a *non-weighted* option which seemed more robust in handling the data from GIS. A prompt from *mlcds* allows a choice between the two weighting options in the code. An illustration of fitting the fixed patterns is given in sec.5.3.2.

5.3 Background on the Observations and Problems Associated with the Analysis Methods

5.3.1 The Normal Incidence Spectrometer Spectral Atlas Data

The Coronal Helium Abundance Spacelab Experiment (CHASE) instrument consisted of a grazing incidence telescope and spectrometer with a spectral resolution of $0.7\AA$, in first order, within the spectral range 160 to $1344\AA$. The instrument had 15 wavebands which could scan sequentially in 10 to $20\AA$ ranges by rotating the grating mechanism through an angle. The instruments' entrance slit corresponded to a spatial area on the Sun's disc of 15" by 180" arcsec.

The data used for this initial study was taken from a series of observations given the acronym NISAT_S. This study is run fairly frequently by the CDS team and by early June about 50 datasets were already available. We required about 10 sound datasets and so 20 were chosen for investigation (accepting that there could be cause for rejection). The details of the observing sequence are given by Harrison & Fludra (1995) but the main points are as follows. NISAT_S only utilises the Normal Incidence spectrometer and its purpose is to obtain a spectral atlas covering the entire wavelength range of the instrument. Atlases are recorded NIS windows 1 and 2. The observations use the 2" by 240" arcsec slit which is moved to the immediately adjacent area on the quiet sun after each 50 second exposure. The slit is stepped 10 times so scanning a total spatial area of 20" by 240" arcsecs in a total observing time of 8 minutes and 20 seconds.

In order to be able to duplicate the method of analysis used in CHASE, the CDS data was organised to match, as closely as possible, the area picked out on the Sun by the size of the CHASE slit. Therefore, the datasets were first summed over half the rasters and then over half the slit with the final data segments corresponding to 10" by 120" arcsecs. This is approximately the same as that observed by CHASE.

Since we wish to apply the differential emission measure technique to quiet sun locations, we want to examine the variability of spectrum line intensities only in such regions. Therefore, we analysed the created data segments in an attempt to identify any presence of active regions in our field of view. The procedure was to sum the data arrays over the dimension corresponding to wavelength and then to examine the variation in line intensity along the length of the slit. The reasoning here was that active areas would reveal a much more pronounced variation from location to location than their inactive counterparts and so be detectable. There was a separate problem with data signal transfer from the spacecraft whereby signal dropouts occurred at single pixels. As an example, fig.5.4 shows the intensity variation with pixel no. along the slit for one of the NISAT_S observations. Note the sudden single point drop of approximately 18.5% between position 30 and 40. Our method for identifying active regions effectively pointed out these signal dropouts too. This technique also reveals intensity spikes due to cosmic ray hits (note the spikes at positions 10 and 80 for example). Standard CDS software is available for removal of cosmic ray spikes but not for signal count replacement. As a result, all the data segments that showed signal dropouts were rejected from the analysis.

During analysis we noticed that there were about 15 redundant pixels at each end of the slit. These were removed from the chosen datasets. Also, we noticed that the spectra were not square to the detector. This can be corrected for using standard CDS software (*vds_rotate*).

Our objectivity in rejecting regions that apparently show active contributions is questionable. Fig.5.5- 5.7 show surface plots of intensity variation with space and time. The x-axis is the spatial dimension corresponding to pixels along the slit (where the edge pixels have been removed) while the y-axis is raster number. The data is from file no. *s3436r00* so there are 9 rasters of 100 second exposures and cosmic ray spikes have been 'cleaned'. The z-axis is the total counts integrated under the line profiles. Note that a photometric calibration was not used at this time. The chosen lines are HeI 515.617$\AA$ (fig.5.5), CaX 574.01$\AA$ (fig.5.6) and OV 629.732$\AA$ (fig.5.7) which cover formation temperatures of $\sim 1.2 \times 10^4 K$, $\sim 1.6 \times 10^6 K$ and $\sim 2.5 \times 10^5 K$ respectively. These correspond, roughly, to the temperatures of the chromosphere, corona and transition zone. The lower and higher temperature regions appear to have a common appearance. The spatial and temporal structure and count ranges are fairly similar. However, the OV line varies much more widely than the other two.

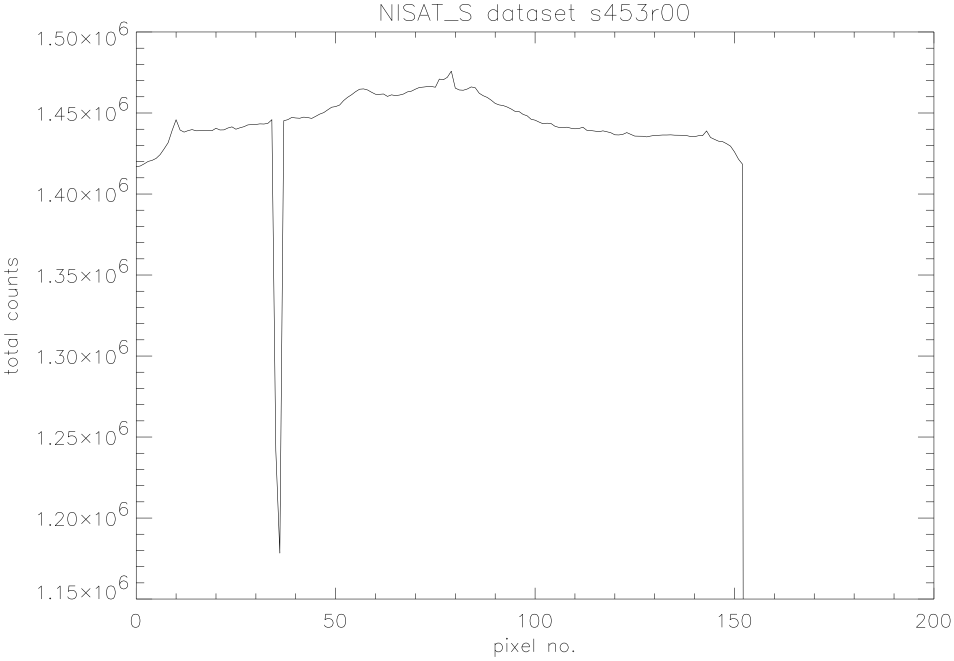


Figure 5.4: Intensity variation with pixel position in the y-direction along the slit. The data is from the NIS2 spectral atlas observations. Rasters 5 to 9 were extracted and the wavelength dimension summed over to produce this plot.

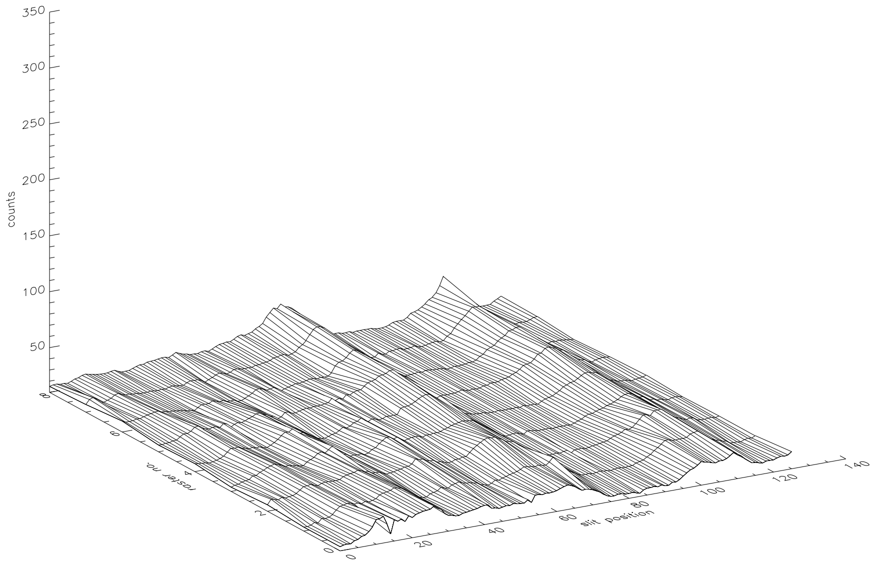


Figure 5.5: Spatial and temporal variation of HeI 515.617$\mathring{A}$ line emission. The x-coordinate is the spatial dimension along the y-direction of the slit. The y-coordinate is the temporal dimension denoted by raster number. The z-dimension is the number of counts.

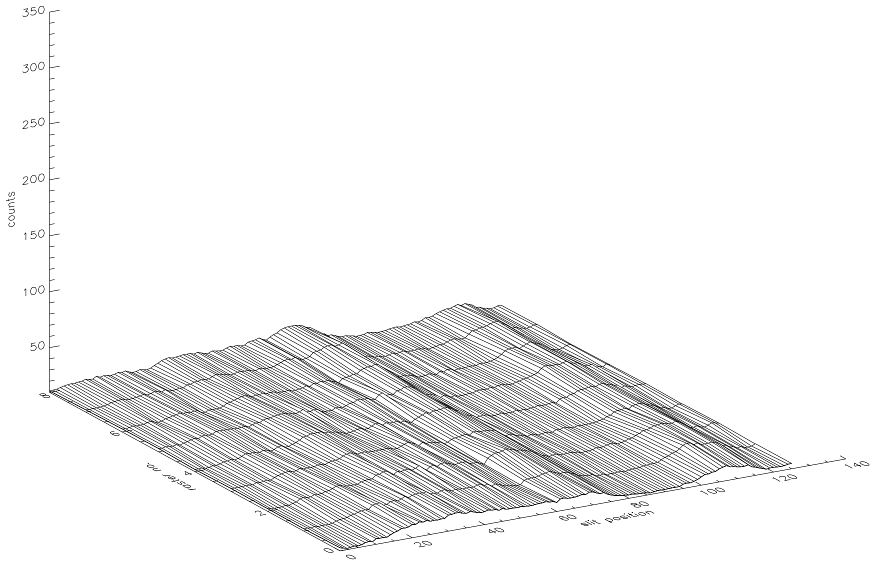


Figure 5.6: Spatial and temporal variation of CaX 574.01$\mathring{A}$ line emission. The x-coordinate is the spatial dimension along the y-direction of the slit. The y-coordinate is the temporal dimension denoted by raster number. The z-dimension is the number of counts.

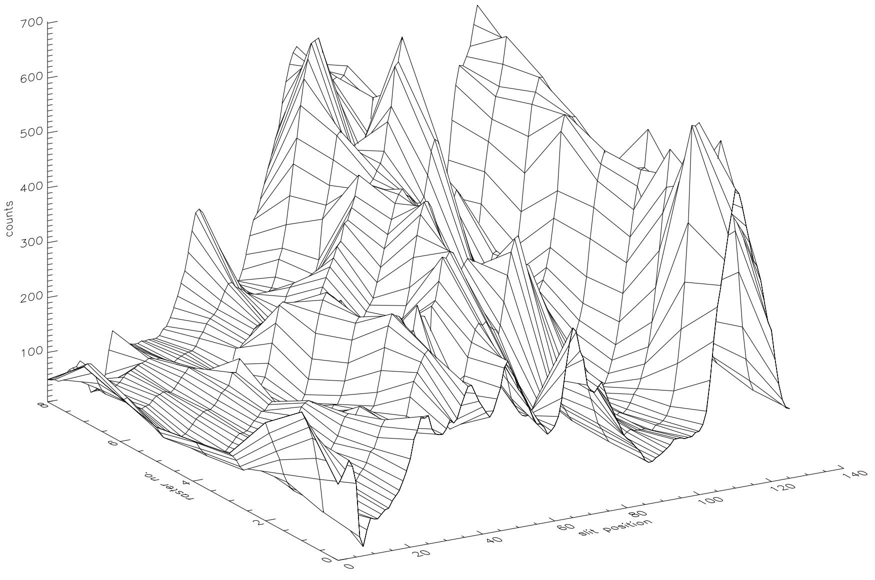


Figure 5.7: Spatial and temporal variation of OV 629.732$\mathring{A}$ line emission. The x-coordinate is the spatial dimension along the y-direction of the slit. the y-coordinate is the temporal dimension denoted by raster number. The z-dimension is the number of counts.

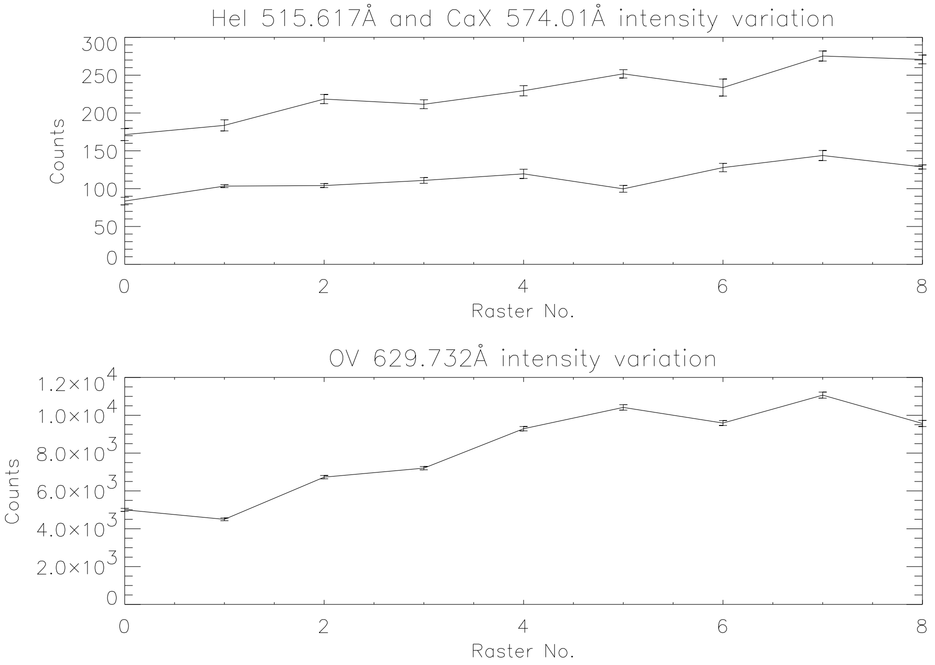


Figure 5.8: Temporal variation of intensities of the HeI, CaX and OV lines. The data is from the NIS2 spectral atlas observations. Rasters 5 to 9 were extracted and the wavelength dimension summed over to produce this plot. In the upper diagram, the HeI variation is the top line and the CaX variation is the bottom line.

Note that the HeI and CaX plots appear on an intensity scale that is half that of the OV plot. In addition, despite some similarities the spatial structure appears more changeable. Fig.5.8 shows the intensity variation with time of each of these three lines at pixel locations 60 to 80 i.e. averaged over that area of the slit. Again, we can see that the OV line varies much more widely than the other two.

Information like this clearly shows the dynamic nature of the solar atmosphere and also shows that intensity variations and ‘hot-spots’ can come and go on time scales close to ionisation relaxation times. Using the estimates for carbon ions presented in sec.3.2.1 we see that both the transition region and the corona are fluctuating over timescales that are too short to have allowed carbon ions to relax to equilibrium. This raises the question as to what is the quiet sun? In the CHASE pattern analysis in

Position	Dataset	Set of Rasters	Portion of Slit
1	s453r00	B	2
2	s454r00	B	2
3	s505r00	A	1
4	s505r00	B	2
5	s633r00	B	2
6	s747r00	A	1
7	s747r00	B	1
8	s747r00	A	2
9	s747r00	B	2

Table 5.4: The final list of NISAT_S observations selected. The letter A refers to rasters 0-4 and B to rasters 5-9. Index 1 refers to the lower portion of the slit with the ends removed (i.e. pixels 15-77) and index 2 refers to the upper portion (i.e. pixels 78-140).

sec.5.5.1 a number of spatial locations are relatively bright but our selection criterion (i.e. less than a factor 3 variation about the mean) accepts them as quiet sun. A number of such files had been marked as active regions by the observer. We rejected these datasets on this basis but decided that it was statistically unsound to reject others of this type. The purpose of our analysis is to quantify the variation of intensities in the quiet solar atmosphere and we did not wish to be biased.

The final list of datasets and segments used is given in table.5.4 along with the letter (a or b) to identify whether the segment corresponds to the first five or second five raster steps in the sequence and the index number (1 or 2) to specify whether the data comes from the lower or upper part of the slit. All the data were stored in IDL save/restore format.

The Earth's magnetosphere normally acts to shield spacecraft from direct bombardment by cosmic rays. SOHO's position at the L1 Lagrangian point is outside the influence of the magnetosphere so it is more susceptible to cosmic rays than Earth orbiting satellites such as YOHKOH. Thus cosmic ray detector activation is a major problem, in particular for the normal incidence spectrometer, which must be addressed. Fig.5.9 shows a normal incidence spectrum from spectral band 1. Cosmic ray events appear as dots and are widespread. We can see that it is not unusual for these events to activate more than one pixel. This presents another problem discussed in section.5.3.3. Note also the appearance of a detector blemish in the lower left portion of the spectrum. In the spectral display mode, the cosmic rays appear as thin bright line-like features. For example, fig.5.10 shows features that are difficult to distinguish from spectrum lines. Note in particular the events at 541 & 593$\AA$. One must seek to remove each of these spurious features as they disturb the spectral line fitting procedure. Fig.5.11 shows the same data segment as fig.5.4 after the removal of cosmic ray features by the CDS IDL code. Note that the peaks at positions 10 & 80 have been smoothed. This cosmic ray cleaning code was applied to all the datasets although its effectiveness was dependent on the type of observation that had been made, so it was not always applied in this form (see section.5.3.3). All the data were stored in IDL save/restore format to be compatible with the CDS analysis software.

As mentioned, the NISAT_S observations were originally intended to provide a spectral atlas for the normal incidence spectrometer. For our work an atlas is very helpful for identifying lines that we can use for the analysis of statistical variance. In fact the atlas had not been fully assembled and so as part of this work we completed a spectral atlas for NIS1 and NIS2 including the cosmic ray cleaning. Many spectral lines were identified and a list of those chosen for the study is given in section.5.5.1 table.5.7.

The line fitting code, *mlcds*, was then used to fit Gaussian profiles to each of the selected lines in each of the selected data segments.

5.3.2 Integration of the Grazing Incidence Spectrometer to the Analysis and Refinement of Observing Sequences

Results of the 'CHASE type' analysis for NISAT_S data are presented in sec.5.5.1. Although the NISAT_S data proved suitable, it does not correspond exactly to the pattern of observations made with CHASE. In the CHASE experiment, the first observing sequence moved the spectrometer over eight adjacent positions across the limb, then returned it to the original viewing position. This raster was repeated a

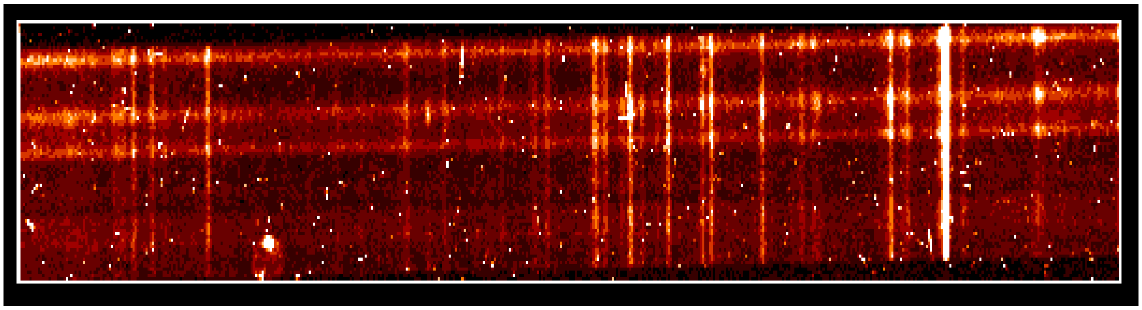


Figure 5.9: Example of how cosmic ray events appear on a typical NIS spectrum. The data is from NIS1 and from the DEMST_1 observations fits file no. s4878r01. The cosmic rays appear as dots and small streaks all over the spectrum.

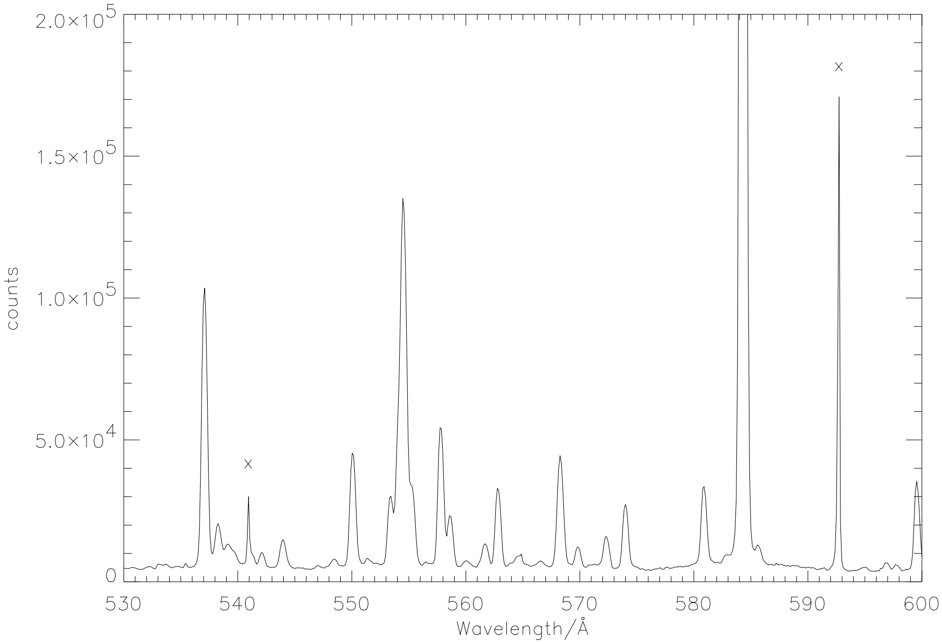


Figure 5.10: Example of how cosmic ray events appear on a typical NIS spectrum. The data is from NIS2 and from the NISAT_S observations fits file no. s350r00. Two cosmic rays are prominent at ∼ 541Åand ∼ 593Åand are marked with crosses.

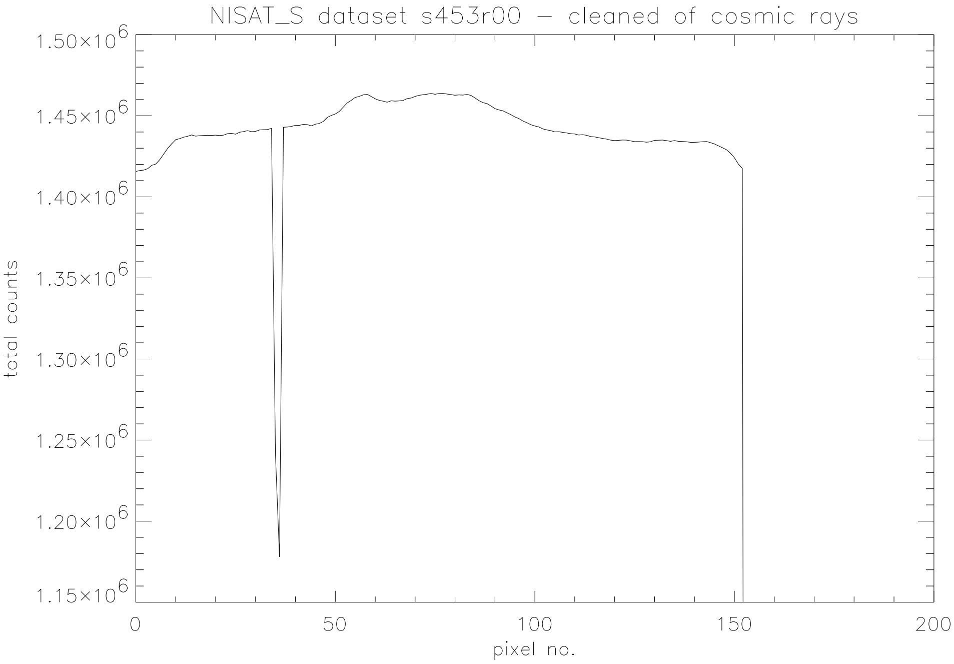


Figure 5.11: Intensity variation with pixel position in the y-direction along the slit after the removal of cosmic rays.

further 3 times. In the second study eight widely separated positions over the solar disc were selected and observed in turn. The purpose of these studies was to observe the temporal variation in intensity coming from the same region of the Sun and then to examine the variation of intensity between different positions. The NISAT_S data was similar to the second CHASE study. We analysed these data but then proceeded to a dedicated study of the temporal variations in a new set of observations. We gave these the acronym RWPM_2. Our main objective was not simply to repeat the CHASE analysis but to establish the statistical uncertainty of the line intensities for the DEM study. Spectrum lines from the Grazing Incidence Spectrometer wavelength region are essential for DEM studies and so the new observations encompassed both NIS and GIS. NIS and GIS cannot observe simultaneously and so the distinct set of observations made for GIS are given the acronym RWPM_1. Both observations were run on 1996 July 7 and were specifically designed to match characteristics of the CHASE ones e.g. rastered area and overall count rates.

Three rasters were made by each spectrometer and these were interleaved to ensure that variations in the solar atmosphere were picked up by both. The GIS observations used the 4" by 4" slit and this was stepped in the NIS slit direction 60 times to cover a total area of 4" by 240". This corresponds to the size of the NIS slit number 5 which was subsequently used for NIS observations. Each exposure of the GIS detector lasted 60 seconds and so the observations took 1 hour to complete. We then switched to the NIS spectrometer whose detector was exposed 9 times for 100s at the same location (allowance was made for solar rotation). This took a total of 15 minutes. The observations were repeated twice more and the complete study lasted a little over 7 hours. The Grazing Incidence Spectrometer observations provided us with the first opportunity to look into the complexities associated with it and also to plan for possible use of the long slit with GIS to precisely match CHASE.

In sec.5.5.2 a complete analysis of the observational accuracy and presence of patterns is given for the NISAT_S data. However, the finalised observations that precisely match those of CHASE are those given the acronym GISAT_S. This sequence was designed using the experience of the data from RWPM_1 and the GIS detector problems discussed in sec.5.3.2 and sec.5.3.2. In addition, new NISAT_S observations are being run concurrently with the GISAT_S ones and pointed at the same location. The main development with GISAT_S is that it utilises the long slit (i.e. 2" × 240"). Therefore, it exactly matches NISAT_S by rastering 10 times in the x-direction. A problem arises. To match the spectral segments used for NISAT_S we require to sum the GISAT_S data over half the rasters. Since GIS is astigmatic there is no spatial dimension to split these over and we do not wish to split the wavelength range. Therefore, the final spatial area covered will be 10" by 240" which is double the length of the NISAT_S segments. Also, it becomes extremely difficult to determine whether data dropouts exist as our previous method of totalling over wavelength effectively reduces the data to one point for each raster. One possibility might be to look at adjacent rasters and try to estimate whether one region contained manifestly less counts than another. However, since the dropouts may be as little as 18% (see fig.5.4) it is possible that they fall within the normal variation between rasters due to the quiet sun.

It was necessary also to design the DEM observing sequence. This was loosely based on the sequences used before. CDS rastered over a 40" by 240" area using the 2" by 240" slit and moving in 2" steps to cover the 40". In addition, it alternated between the normal incidence and grazing incidence spectrometers. At each position the detector was exposed for 200s. This sequence was designed in conjunction with the SUMER team who performed observations covering a 300" by 3" area with the 300" by 1" slit. The CDS part encompassed the same region as SUMER saw and also allowed for errors in the co-pointing. The sequence was named DEMST_1.

Investigation of Detector Anomalies and Aberrations

There are many problems associated with detector response, blemishes etc. which were taken care of and need attention in all observations. However, there are two main difficulties presented by the GIS detector system, namely,

1. The appearance of ghosted lines in the spectrum.
2. Fixed patterning in the line profiles.

Information about the GIS detectors is given by, for example, Breeveld et al.(1992), Harrison et al. (1995) and Kent et al.(1997). Here, we briefly outline the causes of the problems. The detector consists of a stack of micro channel plates (MCP) which produce a charge cloud of electrons when receiving a single photon. The charge cloud varies in number depending on the position of the incident photon. An EUV photon produces $\sim 4 \times 10^7$ electrons and this is the gain of the event. This gain has a normal distribution called the pulse height distribution (PHD). The cloud passes onto a spiral anode detector (SPAN) which consists of strips containing three electrode elements. Digital signals are produced from two of the electrode elements by an analogue to digital converter (ADC). The third element is used as a reference. The magnitude of the first two signals is dependent on the position of the photon event. An on-board look up table (LUT) is used to determine the corresponding pixel position. Separate LUT's are used for different observations dependent on the expected count rates. Raw anode signals are shipped to the ground and fitted by a data spiral pattern to select parameters necessary for re-building the LUT correctly. The spiral pattern takes account of flat fielding effects etc.

The SPAN electrodes are in the form of damped sine waves. Consequently, a large charge cloud covers more 'waves' in the anode pattern than a small cloud. This has the effect of introducing a damped sine wave fluctuation to the spiral pattern. As a result, positions determined on the spiral overlap in some areas and the uniqueness of the solution is lost. This results in transference of data to other areas of the spectrum by movement of lines between spiral arms. Thus spectral lines have ghost companions at different wavelengths and the determination of line positions is affected. The intensity and location of the 'ghosts' varies from observation to observation depending on the voltage and position of the photon event as these affect the size of the charge cloud. However, they can be determined from a knowledge of the detector LUT set up. Software to deal with ghosts is being developed by Mullard Space Science Laboratory.

For detectors which digitise the signals before determining the position, sampling of the continuous function is done at integer values. Clearly this undersamples and introduces an error. This produces a 'fixed pattern' in the line profile. That is, counts are shifted between bins producing an excess in some and a deficit in others. However, the total counts are conserved. Fixed patterning can also be seen on YOHKOH data (see Culhane et al.(1991) and Lang et al.(1992)) but is much less pronounced than that seen with the GIS. We decided to investigate the degree and structure of fixed patterning in the RWPM_1 observations to determine whether we were able to perform spectral fitting in a meaningful way.

In principle the pattern should remain the same for all observations which use the same look up table (LUT) and have similar count rates (Breeveld(1996) - private communication). We decided to check this for our observations since the validity of the LUT may degrade with time (Pike(1996) - private communication). Also, we sought to verify how 'fixed' was the nature of the pattern. Does it vary from line to line? Does a particular line exhibit the same characteristics if we look at different spatial locations? If the pattern is not random, we might hope to reverse the process and recover the original profile.

We extracted the GIS1 data from the RWPM_1 observations. These correspond to the FITS file numbers s3435r00, s3437r00 and s3439r00. The header information in each FITS file contains a parameter GSET_ID which records which LUT was used. For our three sets of observation the GSET_ID was 42. The data arrays extracted spanned three dimensions, wavelength, raster dimension (x) and the 60 movements in the y-direction covering the area of the NIS slit.

Firstly, we looked at the pattern of the strongest line at pixel number 610 along the wavelength dimension and compared it over four different exposures in the y-direction along the slit. The results are shown in fig.5.12. This line was tentatively identified as FeIX 171Å. We see that although there is some variation in the pattern, it is constant enough between the different spatial locations to warrant more detailed investigation. We then split the data into two portions spanning half of the y dimension (slit) and averaged over the x dimension (wavelength). These totalled patterns showed a greater similarity. Hence, despite the variability of the individual pattern we were encouraged to believe that the 'total' patterns would have fixed characteristics. To confirm this we averaged the data over both dimensions, of the slit, for each of the RWPM_1 datasets and found that the patterns matched very closely. Fig.5.13 shows the similarity of the totalled patterns of individual lines in each of the three datasets.

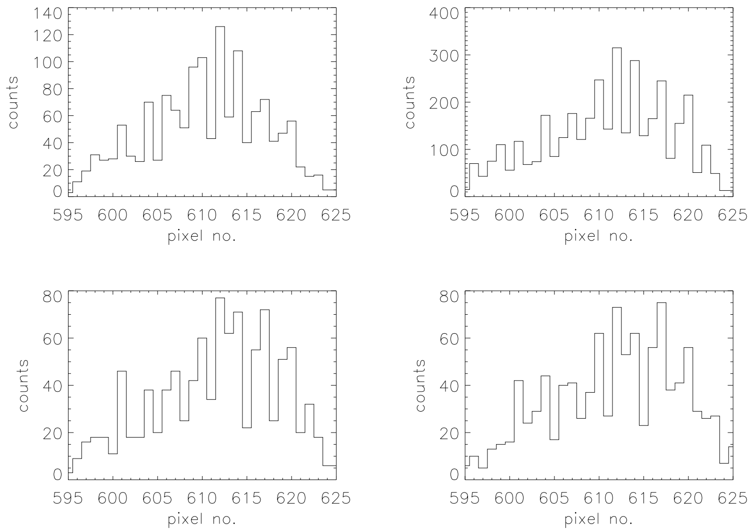


Figure 5.12: Plots of a spectral line at 610 pixels (in the wavelength dimension) which was tentatively identified as FeIX 171Å. The x-coordinates are pixel numbers along the wavelength dimension. The y-coordinates are recorded counts. The four graphs show the pattern of the same line varying at four randomly selected unknown y-pixel positions along the slit dimension.

This is what we would expect since the LUT numbers were the same for each dataset. Also, since the observations spanned 7 hours, we can conclude that any degradation in the applicability of the LUT's is not important over the period of these observations.

Note also that the two lines in fig.5.13 show different fixed patterns. This is consistent with the variability we found over spatial locations even for a single line. However, this does not appear to have much influence on our analysis. The main concern was that we would be unable to say conclusively that the same LUT's give us the same patterns but it appears from fig.5.13 that they do. Differences between individual lines are actually of less concern as the main purpose was to see if it was feasible to apply some retrospective-correction to each line profile.

Each position on the spiral pattern corresponds to a value of the event gain function. Therefore, the sampling grid needs to be fine enough to map every point. After discussions with Pike & Lang (1996 - private communication) it became clear that the resolution of the on-board sampling grid was the main problem and a retrospective-correction was not possible. That is, the only proper solution would be to attempt to obtain the raw data from the spacecraft and apply a finer sampling grid ourselves to reproduce the gain function and obtain the photon position. This is impracticable as the data telemetry rate is such that to get a reasonable signal to noise ratio the transfer time would have to be two to three orders of magnitude longer. Also, the normal planning and on-board pre-analysis software would have to be bypassed. In any event, for this type of detector, there is a fundamental limitation to the resolution obtainable. For a particular charge, falling on the anode, the digitised results will differ by an absolute minimum of 1/charge (Geesman et al.(1991)). Therefore, it is necessary to handle the patterns as they are.

Finally, we should mention one more restriction. The GIS to command and data-handling system (CDHS) link only operates up to a maximum constant rate of 8.9×10^4 counts/s. However, the detector readout electronics can operate up to 1×10^5 counts/s. Therefore, it is possible that count rates above the capabilities of the CDHS are sent to it. If this happens, half the data from the detectors are lost. In addition, above 5×10^4 counts/s the actual total count rate is unknown (Kent et al.(1997)).

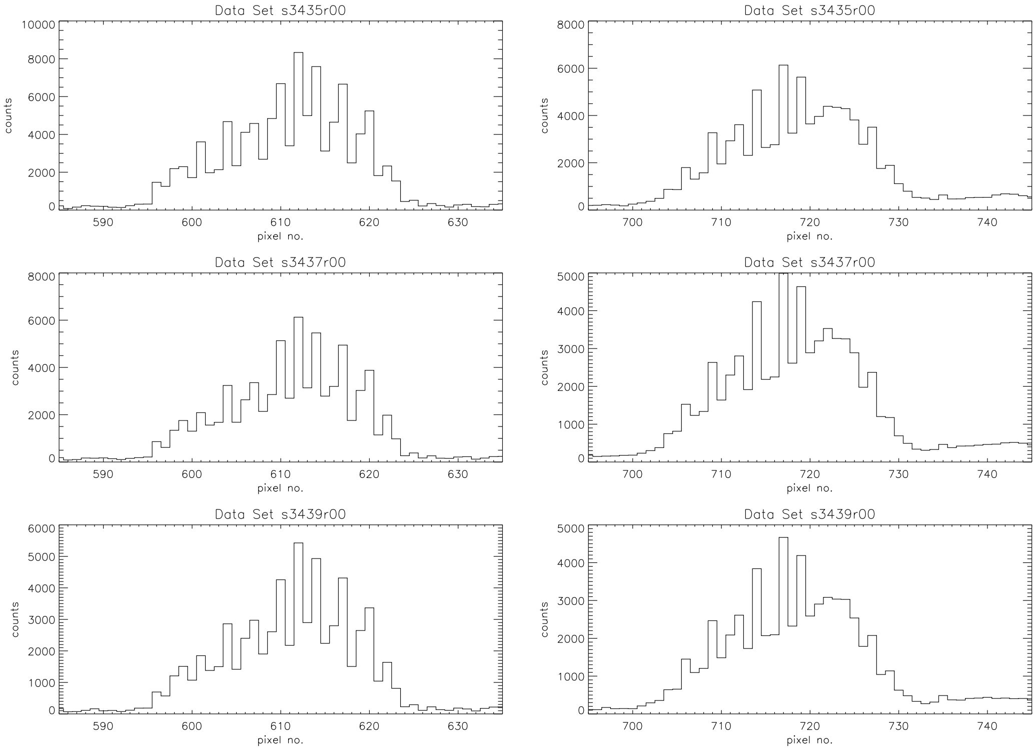


Figure 5.13: Example showing the similarity of individual lines' patterns when the counts from all 3 datasets have been totalled. Note also the pattern variation between the two different lines. The x-coordinates are pixel numbers along the wavelength dimension. The y-coordinates are recorded counts. The line in the 3 left hand diagrams is the line previously identified as FeIX 171Å(see text). The line in the 3 right hand diagrams was unidentified.

Test of Gaussian Line Profile Fitting to Fixed Patterns

We determined their wavelengths using the available wavelength calibration in the CDS software (*pix2wave*). Then, we attempted to find the options in *mlcds* that would allow us to get the optimum fit, regardless of whether the fit was good. Hence, in the results that follow, the same options have not necessarily been used for each pattern and, in some cases, two lines have been fitted instead of one. Evidently, handling the GIS lines is more time intensive than NIS and the development of automatic procedures is not necessarily of benefit here. As discussed in sec.5.2.3 *equal weighting* was introduced solely for this purpose, therefore, results are presented for each chosen line with and without this option. The total flux count from the line profile fitting routine was compared to a hand count. The results are given in table.5.5. We selected 8 lines spanning the four GIS detectors which showed different types of pattern.

It is clear that the program's ability to handle the data is determined by the pattern type and the fit options chosen. Fig.5.14 shows an example of attempts to fit two types of pattern using the intensity weighted and equal weighted fits. We can see that in the first pattern both options fit the data extremely well. The comparison to the hand count, from table.5.5, shows that both overestimate the total counts slightly, 2.8% for the weighted fit and 0.1% for the non-weighted fit. In contrast, the weighted fit fails on the second pattern underestimating the total counts by 78.3%. It appears that when the pattern is almost random the intensity weighted fit works quite well, but that when there are large single pixel drops in the profile, the code attempts to find more than one line. On the other hand, the fit with equal weighting is within 0.8%.

In the majority of the weighted fits the errors in comparison with the hand count are within the estimated standard errors of the program and in the equal weighted cases this is always true. Also, the equal weighting fit is generally within 1% of the hand count. In fact, we only found one instance where this was not the case i.e. the line at 788.2Å. In general, the equal weighting fits were very good with the weighted fits not quite so good. We conclude that with judicious selection of lines, based on their patterns, we can fit the total counts in each to within our attempted statistical accuracy of 1%. Also, if necessary it is possible to take those lines that

Approx. Line λ (angstroms)	ESEF%	Comparison to hand count	Weighted Fit	Back-ground
171.0	9.6	2.8% +	Y	sloped
171.0	11.1	0.1% +	N	sloped
174.5	8.0	8.4% -	Y	flat
174.5	10.8	0.2% +	N	flat
188.2	51.9 & 51.8*	78.3% -	Y	flat
188.2	17.6	0.8% -	N	flat
305.3	20.1 & 11.0*	18.6% -	Y	sloped
305.3	27.6 & 11.0*	0.1% -	N	flat
393.5	17.2	12.9% -	Y	flat
393.5	14.5	0.7% -	N	flat
465.2	10.8	15.1% -	Y	flat
465.2	10.5	0.7% -	N	flat
770.4	13.1	5.1% -	Y	flat
770.4	15.3	0.8% -	N	flat
738.2	12.3	6.7% -	Y	flat
738.2	12.6	3.1% -	N	flat

Table 5.5: Comparison of line fitting options for GIS data. * denotes that two lines have been fitted, +/- denote that the fit was too high/low in comparison to the hand count.

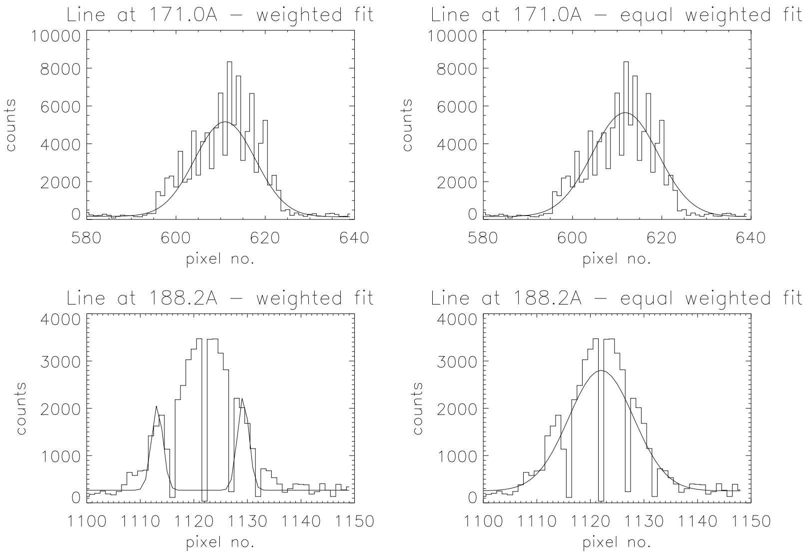


Figure 5.14: Weighted and non-weighted Gaussian line profile fits to the GIS1 lines at 171.0Å and 188.2Å.

cannot be fitted to sufficient accuracy and normalise them to the hand count in order to achieve the 1% statistics. This study is encouraging for all analyses reliant on intensity information from the GIS. In our case this is of major importance for the DEM study. Finally, we note that no attempt has as yet been made to retrieve line profile information, e.g. doppler components, and the above analysis assumed that in each case the line being dealt with was a single stationary Gaussian or a blend of two.

5.3.3 Modifications to Cosmic Ray Cleaning

The standard CDS-IDL cosmic ray cleaning routine (*avg_wo_cr*) was used successfully on the NISAT_S data. However, it automatically averages the data over one of the dimensions. If the data array consists of wavelength, slit and raster position dimensions then the routine reduces the data array to two dimensions, namely, the wavelength and slit. There is some flexibility here but contraction on wavelength is unsound (see below). The code works by locating cosmic rays as deviations of greater than 3 σ in pixel count along the dimension to be contracted, marking their pixels, and then missing them out when the average is performed. However, it is quite possible, see fig.5.9, that cosmic ray hits may not occur perpendicular to the detector, that is, they leave a track across adjacent pixels. In such cases the routine is less effective and the choice of data dimension that can be worked over to locate them is constrained. In the case of a glancing hit, the slit dimension will not produce the best results (see fig.5.15 where the partially removed hit had smeared over pixels 60 to 89 in the y-direction along the slit) but contracting on the raster dimension means that the adjacent positional information is lost. For the NISAT_S study this was not a problem. We had to average in the raster direction to match the size of the CHASE slit. However, the aim of the RWPM_2 observations was specifically to map the intensity variations in time. As a result, we did not wish to lose time resolution by averaging over this dimension. Another approach was required.

One possibility is to write a new 'cleaning' routine that maintains the original array dimensions of the data. This would work, by finding values that deviate from those around them by more than 3 σ, and replacing them by some form of running average over, perhaps, three neighbouring values. In practice, we decided to make a modification to the existing routine to allow us to separate the data array dimension that was chosen for 'hit' location, from that chosen for averaging. This means that we can locate and mark cosmic ray hits in the most appropriate way for the particular observations and then average the data in a different way. It was felt that the method likely to be most successful was to locate hits in time, reasoning that the likelihood of a cosmic ray hitting exactly the same spot within the time frame of the observations was extremely small. This may not be true for long exposure times. To recap, for the RWPM_2 data we used the new cleaning code, temporal dimension for hit location and the spatial dimension for averaging. The method worked extremely well. Fig.5.15 shows the removal of a cosmic ray that had previously proved too stubborn when the original program was applied. The left hand diagram is a portion of the original spectrum. The centre frame shows the same portion after the original routine was applied and the right hand frame shows the results after application of the modified routine. Note how the hit at pixel no. 258, which was only partially removed by

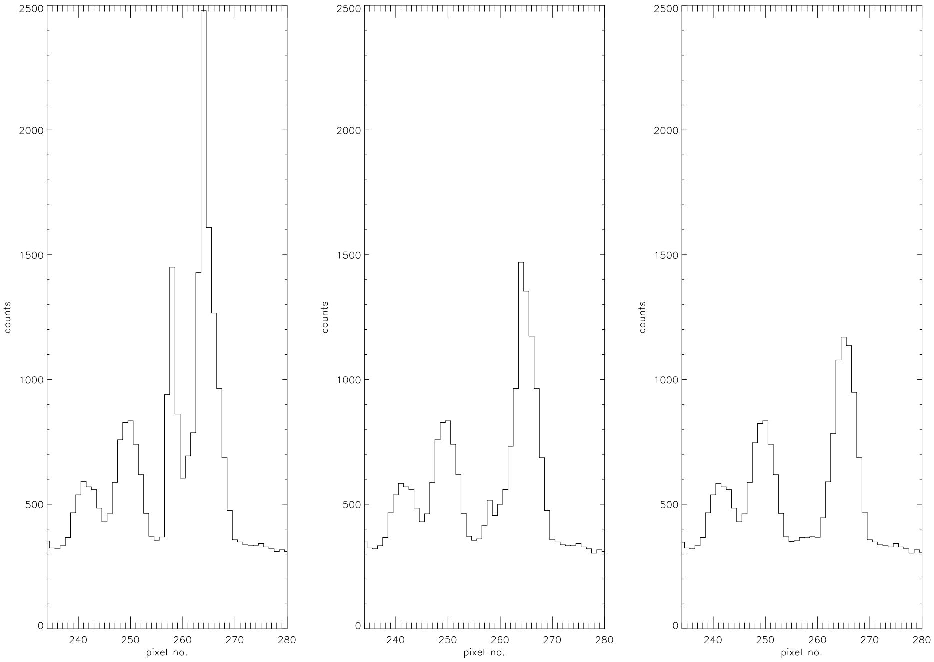


Figure 5.15: Application of cosmic ray removal routine to a hit between pixels 60 and 89 in the y-direction along the slit. The x-coordinate is pixel number along the wavelength dimension. The y-coordinate is recorded counts. Left - original spectral segment. Centre - segment after cleaning along y-direction with original routine. Right - segment after cleaning with modified code by location along teporal dimension and averaging over the y-direction as before.

the original routine, is completely removed by the new one. Note also that the 'real' spectral lines at 242 and 250 have count rates that are unchanged. Interestingly, it appears as though a cosmic ray lies atop the line at pixel no. 264. The routine has also been successful at differentiating it from the line itself. It is clear that this method is suitable for this particular data. Development of an array dimensionality maintaining cleaning code is deferred until the need arises.

As mentioned, for DEMST_1 the data for NIS were interleaved with GIS therefore only individual rasters were present in the datasets. More precisely, the data arrays were of the dimensions (1024,1,160) i.e. wavelength, raster, slit. We proceeded to test the cosmic ray cleaning on both the wavelength and slit dimension. Fig.5.16 shows

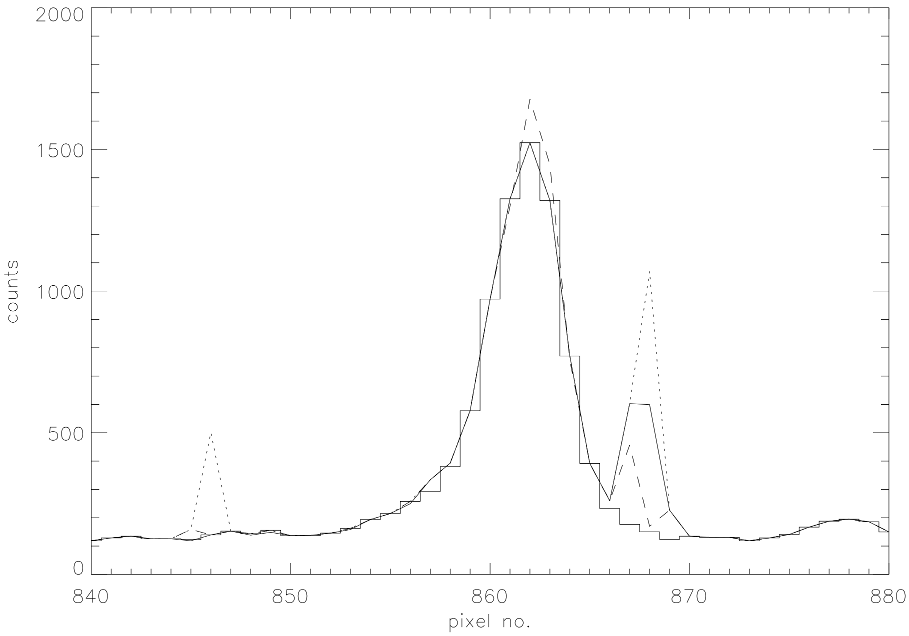


Figure 5.16: Example of improved cleaning as a result of collecting the DEMST_1 data together. Histogram - preferred method of cleaning over rasters. Dotted line - original data. Solid line - cleaned along slit. Dashed line - cleaned along wavelength. These are shown to highlight their inability to remove the event completely. Our method maintains the data and removes the spikes.

NIS1 data for fits file no. s4878r01. The data was extracted from pixels.840→880 along the wavelength dimension and 30→40 along the slit. The dotted line shows the original data summed over the slit dimension. Note the spikes at pixels 846 & 868 due to cosmic ray events. The solid line shows the result after locating hits on the spatial dimension. The dashed line shows the result after locating hits on the wavelength dimension. In both cases the data were averaged over the slit to produce the plot. We conclude that spatial location is a better method than wavelength location. Note that both methods almost completely remove the spike at pixel no.846, but that neither method has successfully removed the one at pixel no.868. Note also that locating by wavelength has added counts atop the spectral line at pixel no.862 and been marginally less effective in removing the line at pixel no.846. This method is clearly unsound. Given the success of the method used on the RWPM_1 data, despite appearing better, location spatially is unsatisfactory. We decided that location by comparing rasters would be more effective. Thus it was necessary to develop a procedure to read each dataset individually and stack the data arrays together. Then the cosmic ray location could be done over the raster dimension and the averaging performed over the slit. This retains the spatial components of the data. For our purposes we split the slit in two rejecting the redundant pixels at the ends as before. The method proved satisfactory. The histogram overplotted in fig.5.16 shows the result after separation of the stacked data.

5.4 Outline of Analysis Procedures and Preparation for the DEM

5.4.1 Coupling of the Observational Data to the DEM Code

Following the benchmark study of DEM programs by Harrison & Thompson (1992) it was agreed to adopt the inversion routine of Brown & Thompson (1991) as the standard preferred method. The core of the routine has been restructured and interfaced with ADAS, as routine ADAS601, by Summers et al.(1996). It has been thoroughly tested and validated on analysis of SERTS data (Lanzafame et al.(1997)). Following discussions with the CDS team, Summers et al.(1996) also prepared a specific format for intensity files for entry to the DEM code. It is our wish to route the spectrally analysed, corrected and calibrated experimental data directly into files of this format.

The experience gained, during the work presented in the previous sections, has enabled us to develop a comprehensive analysis procedure and data archiving system that is both compatible with the CDS data and software and allows easy conversion to the ADAS DEM format files. In this section, the analysis path, and descriptions of software developments, are presented in detail.

The routines outlined below have been written to be flexible and applicable to different types of observing sequence. However, variety in all possible observing sequences is manifestly difficult for the *clean_data* routine to handle. For example, in the case of opacity or dynamic event studies, individual lines require to be analysed and mass automatic production of fits to them is not appropriate. Rather one wishes to fix certain lines and allow the positions of doppler shifted components to move. Therefore, although *clean_data* applies many important and necessary corrections to the data, it is not an all encompassing routine. Minor modifications may need to be made to accommodate individual observing sequences. The possibility of generalising it for all observing sequence data does arise through the standard format of the detector data arrays. All the data are collected in an array of the pattern; (wavelength, X, Y, time). However, gross averaging and reduction of data is not always necessary as exposure times may be such that good statistics can be obtained with small slit and raster sections. The principal point is that since it is not possible to account for or foresee every type of data, the initial data entry formats should not be strictly prescribed. Only after preliminary individual manipulations are the tighter data structures mandatory. In the case of the CHASE type analysis and DEM, the method and data use has been fairly well defined so it is possible to tighten and automate the initial manipulations.

Finally, all following routines prepare data for the line fitting program and store the results in a standard format so that any subsequent analysis procedures can easily be applied to it all. In addition, archiving of header information and identification tags is forced at each stage to allow the route to source data to be easily traceable.

5.4.2 Overview of the Analysis Path

Fig.5.17 is a flow chart showing the analysis path and archiving procedure. The manipulations discussed in the previous sections have been reduced to a number of small routines which can, if necessary, be tailored individually to any new developments. The routines are written in IDL and make use of a number of procedures provided by the CDS-IDL software. Therefore, the commands are typed interactively at the CDS software IDL prompt.

After starting *cidl* the routine *clean_data.pro* is invoked to separate selected data into segments suitable for the particular analysis. In the case of the analysis of statistical variance this refers to the spectral segments summed over half rasters and half slits corresponding to the size of the original CHASE slit. In the case of the DEM study, the spectral segments are untouched for GIS. For NIS the data are split along the slit, collected together, cleaned for cosmic rays on the time dimension and finally separated in time again (see sec.5.3.3). First, a CDS data fits file is selected by response to the prompt *'Enter data set for analysis'*. The chosen file is then read in by the standard *readcdsfits* routine. The program then checks, internally, to see whether the data was recorded by the normal incidence or grazing incidence spectrometer. In the case of the former, a number of standard corrections are performed to account for effects inherent in the detector. In the case of the latter, methods have been developed to deal with the data stream and investigation of the spectrum characteristics has been made (sec.5.3.3). Initial required corrections however are less well understood at this time and will be incorporated at a later data as they are not essential for the current study and demonstration which focuses on NIS.

For NIS the data has the *debias* correction imposed. The transmitted data will contain photon counts due to spectrum line radiation, continuum radiation, and possibly scattered light. In addition, there will also be a 'dark' current noise associated with the detector itself. However, the Viewfinder Detector Subsystem (VDS) CCD, of the NIS, adds a thermal background to the signal transmitted to the ground. This is because the temperature of the instrument is higher than was originally anticipated. This *bias* needs to be removed before analysing the spectrum (in extreme cases the bias can swamp the entire spectral signal (Harrison et al.(1995))). In order to correct

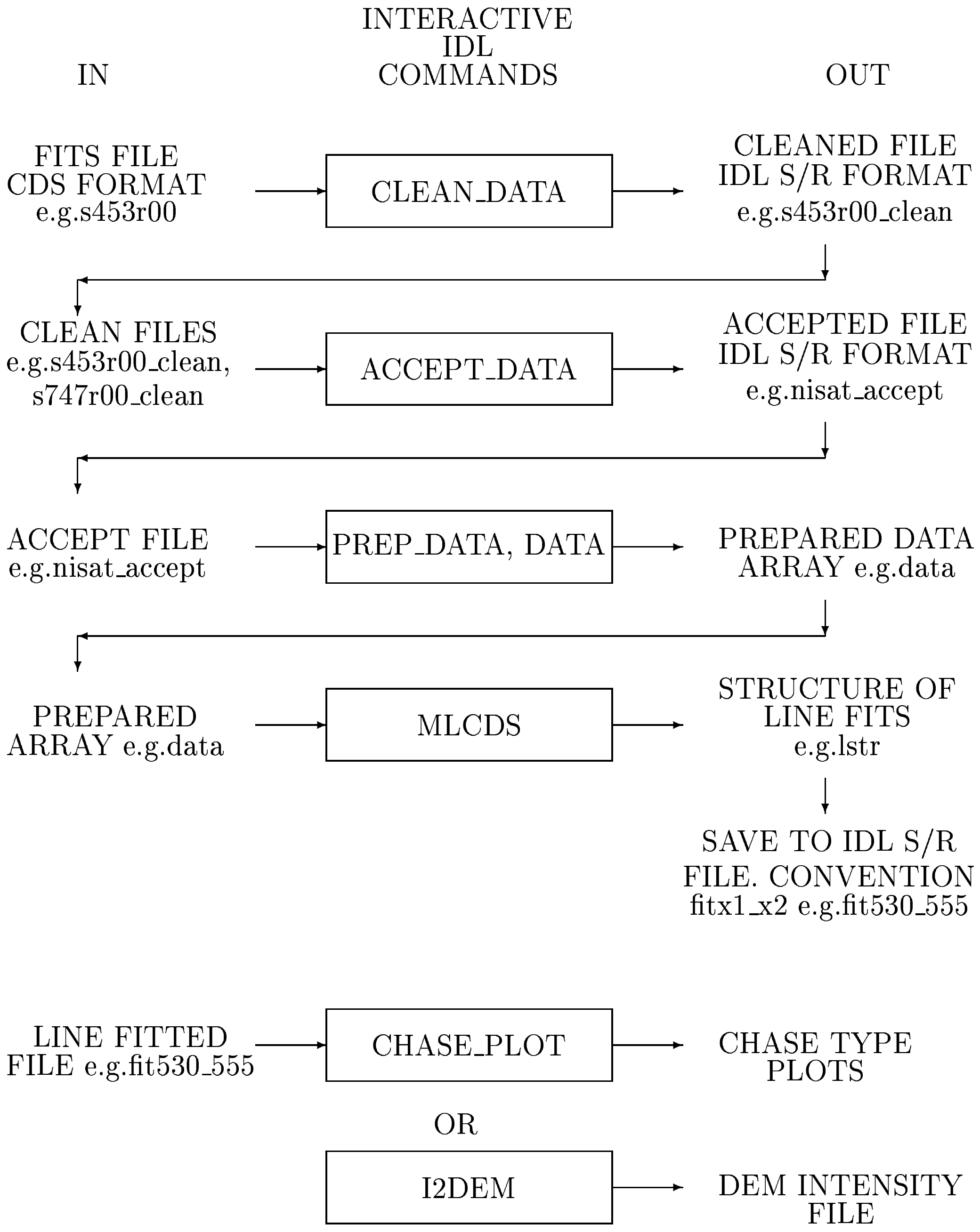


Figure 5.17: Schematic of CDS data analysis path and software components.

for the bias, data is transmitted from the four corners of the detector but outside the area of the two spectra. Thus it should only contain the bias, some scattered light and the inherent 'dark' current noise. It gives a measure of the magnitude of the bias together with an uncertainty, which can be subtracted from the spectral data. A CDS software routine *vds_debias* is available to do this and it is present in *clean_data.pro.*

Second, the routine *vds_calib* is applied to the data to convert it to photon events per pixel per second, rather than ADC's per pixel per sec. This conversion depends on the voltage of the detectors (Thompson (1994)). In addition, the routine divides the data by, and thus corrects for, the flat field image (or an appropriate part of it). It also takes account of any 'burn-in' of the detector due to the consistent appearance of strong lines e.g. HeI 584$\mathring{A}$. Finally, a correction term, to account for the nonlinear response of the detector, is applied if the count rate is high enough. This is introduced at 300 ADC/pixel/sec ($\sim$ 120 photon events/pixel/sec at 756V) but only becomes significant above a certain threshold ($\sim$ 5000 photon events/pixel/sec at all voltages (Thompson(1994) - CDS software note 13)).

For the DEM observing sequence the intensity calibration is then applied using *nis_calib*. This routine converts the data to physical units and for compatibility with ADAS we return the data as photons/cm^2/sec/steradian. The complete pre-flight calibration report is still under discussion so *nis_calib* should be subject to review after its publication.

The data for the two windows is then extracted from the structure and corrected for the fact that the slit images are not parallel to the edge of the detector. In addition, the spectral lines are slightly tilted. This 'skew' is wavelength dependent and varies between NIS1 & NIS2. For NIS1 the variation is approximately linear and averages to $\sim$0.004 pixels per pixel. For NIS2 the variation is approximately quadratic and of the same magnitude. A 0.004 'skew' corresponds to a slant along the spatial dimension of the slit of up to $\sim$0.64 of a pixel, for example. This in turn amounts to $\sim 0.045\mathring{A}$ and $\sim 0.075\mathring{A}$ for NIS1 and NIS2 respectively. A small routine was written to convert from pixel space to tilt factor and this was used to correct data when fitting the line profiles (see below). This correction is made for completeness and high precision though it is likely that, since the data are only binned to one

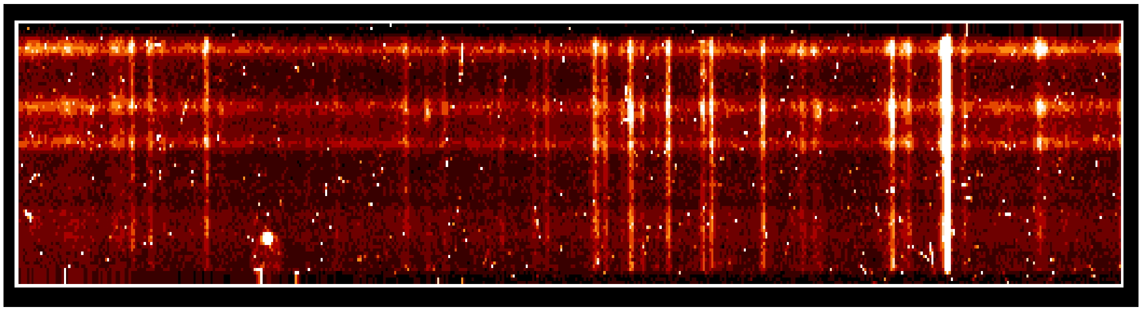


Figure 5.18: Example of the correction for spectrum slant applied to NIS1 dataset s4878r01. Note how the spectral lines now appear parallel to the slit edges and the redundant pixels at the lower edge are more clearly visible.

pixel, it will not improve our ability to separate lines within one pixel of each other. In addition, after testing we found that the difference in the flux in the lines was typically less than 2% and that this was within the estimated standard error of the fit. Fig.5.18 shows the same spectrum for NIS1 as in section.5.3.1 fig.5.9, but after application of the correction software. A comparison between the two shows the effect and the success of the correction.

clean_data then makes a check on the dimensionality of the data arrays to determine whether the data is from the DEM study (acronym DEMST_1) or the spectral atlas surveys (NISAT_S and GISAT_S). The data are then split and averaged as described above. For NIS, cosmic ray cleaning takes place at this point using the modified cosmic ray cleaning procedure described in sec.5.3.3. In addition, care is taken to ensure that missing pixels are accounted for correctly. Inadequate attention here initially caused some datasets to suffer major corruption. Data can sometimes be lost from individual pixels and pixels at the spatial extremes of the spectrum are often redundant (see section.5.3.1). These pixels are marked with a value of -100. This can obviously bias any averaging or totalling especially for NISAT_S where the intensity calibration is not applied and the values are consequently lower. For DEMST_1 the calibration affects the value that the missing pixels are marked with so a check is made to see whether any exist, prior to the calibration, in order to mark the value

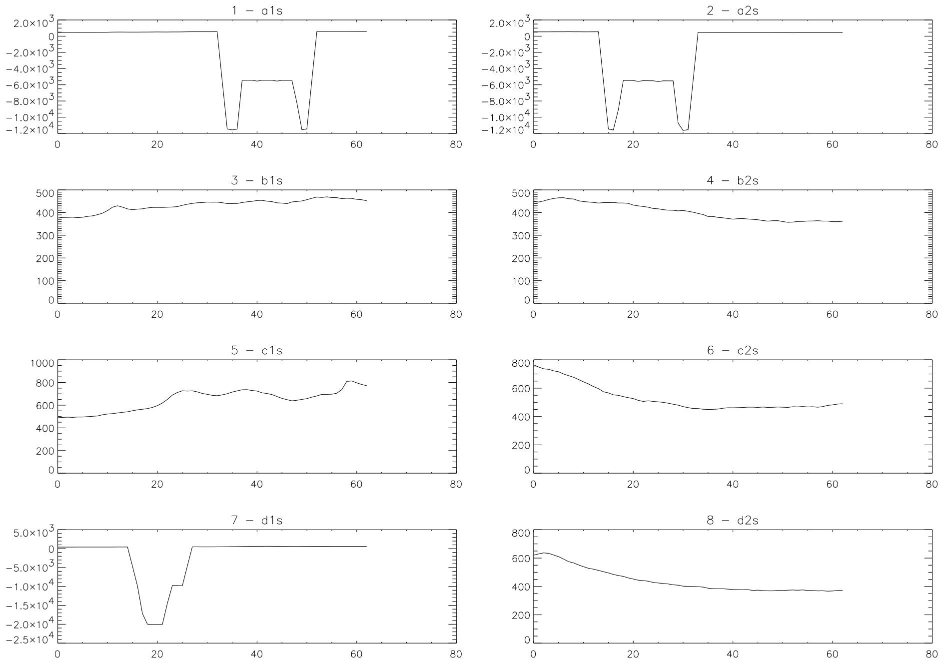


Figure 5.19: Example of the Display for Accepting Data segments. The x-coordinates are the pixel numbers in the y-direction along the spatial dimension of the slit. The y-coordinates are the recorded counts.

correctly. Subsequently, cosmic rays are marked with the same value so that the two features can be 'cleaned' simultaneously. The data are archived in IDL save/restore files with the naming convention tag (_*clean*) appended to the dataset name to distinguish it. Each individual spectral segment is given an index name corresponding to which portion of the rasters, slit and which detector it is. Data segments are also stored which are summed over wavelength, rather than slit, to aid subsequently in identifying signal drop outs. [1] In addition to data segments the header information, about the observing sequence, is also stored in the file.

The next step is to view the data segments to determine whether signal drop outs

---

[1] This works well for NIS but is problematic for particular sets of GIS observations. This is simply because the data arrays have only one dimension, or a wavelength dimension and 10 spatial locations, for the DEMST_1 and GISAT_S observations respectively. In either case, the data are effectively reduced to single values for each (if any) spatial raster.

Index	Detector window	Raster nos.	Slit Pixels
a1s	1	0 → 4	15 → 77
a2s	1	0 → 4	78 → 140
b1s	1	5 → 9	15 → 77
b2s	1	5 → 9	78 → 140
c1s	2	0 → 4	15 → 77
c2s	2	0 → 4	78 → 140
d1s	2	5 → 9	15 → 77
d2s	2	5 → 9	78 → 140

Table 5.6: Partitions of the NISAT_S data into segments.

occur and if there is evidence of other data corruption. This visual inspection is useful in convincing oneself that the cosmic ray cleaning has been performed successfully. A routine *accept_data.pro* is invoked which plots the cleaned spectral segments, summed over the dimension of wavelength dispersion, and allows the user to accept or reject them interactively. A dataset is entered in response to the *'Enter a CLEANED dataset for analysis'* prompt. This must be of the xxx_clean format output from *clean_data.pro*. An example plot for the dataset s453r00_clean is given in fig.5.19. The index numbers correspond, for this observing sequence, to the divisions given in table.5.6. A simple yes/no response to the question *'Accept plot 1 ?'* will stack the accepted plots into an array suitable for entry to the line profile fitting routine. After eight segments are plotted the option to *'Continue and analyse another CLEANED dataset ?'* is given. In this way, the array can be stacked with unlimited data segments. An internal check is made to ensure that, if a data segment is rejected, the corresponding segment for the other detector is also rejected. An identification array is created that stores the dataset name and segment letter designation to allow back tracing to the original header information in the xxx_clean file. Both the data array and identification array are stored in IDL save/restore format files with the naming convention *nis_accept* or *gis_accept* dependent on the spectrometer. In addition, a check is made every time a new dataset is analysed to ensure that the different xxx_clean datasets relate to the same spectrometer.

The data arrays stored in the xxx_accept files can be entered directly to *mlcds*. However, since it can only fit up to ten lines, it is usual to extract some segment of the array that covers only one or at most a few lines. This is aided by the routine *prep_data.pro*. A dataset is entered at the prompt *'Enter an xxx_accept IDL save/restore file :'* and a detector number is also specified at *'Enter a detector number :'*. The line or lines' pixel range can then be extracted by giving the lower pixel limit and then the estimated range in response to the questions *'Enter lower extreme pixel no. of line(s) to be fitted :'* and *Enter total pixel width of line(s) to be fitted :'*. The data array is restored to the *cidl* prompt by passing a variable, e.g. data, as a parameter in the call to *prep_data.pro*. All this requires that the pixel range of the line is known. A keyword */rotate* is available to allow the correction of tilted lines. This is done by the routine *vds_rotate* in conjunction with the routine *tilt_factor*. *tilt_factor* uses the central pixel of the entered range, along with supplied CDS software fits to the tilts of each detector, to calculate the parameters to pass to *vds_rotate*. Subsequently, the prepared data array is entered to *mlcds* which has the variety of options, allowing it to fit Gaussian profiles to the lines, described in sec.5.2.3. A comment is usually entered at this stage (in the header field - see sec.5.2.3) which gives the name of the user, the date of processing and the xxx_accept file that was used. This maintains the traceable route to source. The suggested convention is to give the output of the fits structure the name *lstr* and to save it to an IDL save/restore file, together with the minimum and maximum of the pixel range in the name e.g. fit530_555. This ensures that subsequent analysis routines, such as the 'intensities to DEM file' code and the statistical variance codes, can be used. The output structures are set up so that a series of lines can be stored in one file. It was thought, therefore, misleading to identify the file by one wavelength or pixel number and more appropriate to give a pixel range covering each line in the file. Also, a definitive statement of what the line is was thought impractical, anticipating the results presented in sec.5.5.1, so no attempt is made to introduce this in the naming convention. However, *mlcds* has a facility to make a preliminary identification if desired (see sec.5.2.3).

The structure based IDL save/restore files make the final fitted data compatible with the CDS software and easy to handle in subsequent analysis routines. Two procedures for further application have been completed. Firstly, the routine which creates the CHASE type plots accesses the *fitx1_x2* files and makes use of any preliminary line identification so that it is easy to use. Invoked by typing *chase_plot* a *fitx1_x2* file is entered at the prompt and a *Preliminary Identification* list with an *Index* is displayed on screen. The program then asks the user to *Enter number of line in file for analysis :* whereupon the line fluxes and error are extracted, normalised to the average of the positions, and the plot displayed or written to a file. This routine was used repeatedly and successfully in sec5.5.1.

Finally, the code which writes the DEM intensity format file is started by typing *i2dem.* This routine prompts for the name of an output intensity file (*Enter a DEM file name :*) and then an input archive *fitx1_x2* file (*Enter an archive file name :*). The code determines how many datasets were used in the line profile fitting and displays a message on screen (e.g. *There are 40 datasets in this file*). The fits to the lines are stored in the data structures and so one of these is selected by response to the prompt *Choose a structure element to analyse (0,1,2...).* From this the code extracts any preliminary line identifications to list on screen along with an index number. If no identification is available a *u* is displayed. Selection of a particular line is made by entering the appropriate index number. Further lines can be selected or alternatively lines from different datasets can be added. In this way the intensity file is built up. The line fluxes and estimated standard errors are then extracted. The estimated standard errors are converted from percentages to actual values and both are then written to the specific format of the DEM intensity file defined by Summers et al.(1996). An example is given in sec.5.6.1 fig.5.27. The line identifications can be given in two forms and are parsed to write out in the intensity file. The forms are e.g. 'OV 629' or 'O+4 629'and are entered while fitting the line profiles with *mlcds.* The *observed wavelength* is taken from this string and the fluxes and estimated standard errors are placed in the columns under *observed intensity* and *uncertainty.* A keyword option */wvspect* can be invoked to write the wavelengths into the *spectroscopic wavelengths* column also. This was found useful in practice. Other information has to be entered by hand as it relies on building up the $G(T_e)$ collection file as described in sec.5.6.1.

Note that the data stream is also compatible with the output of the SUMER instrument. The IDL save/restore files that it produces are already in the correct format for *mlcds* and so can be directed to *fitx1_x2* archive files also.

5.5 Variability of Spectral Line Intensities from the CDS Data

5.5.1 Analysis of the Instrumental Accuracy and the Quality of the Observations

In order to obtain useful information regarding the physical nature of the solar atmosphere we must first overcome a number of problems associated with interpretation of the spectral line intensities. It is clear that the useable measurements, obtained by the spectrometer, depend on physical and observational factors. Included in the former category are changes in the structure of the solar atmosphere and in the physical conditions of the emitting plasma, during the time of the observations, along with changes in the intervening regions along the line of sight between the target and the spectrometer. In addition, the spectral reduction depends on the atomic model adopted. In the latter category, errors are introduced by the instrument, the methods of analysis and the instrumental sampling technique. These include random errors associated with the detection of photons, the accuracy of the intensity calibration and the instrumental transmission, the efficiency of the data communication system and transference to the ground and accuracy of the line profile fitting techniques. Also, the spectrometer will only 'see' a proportion of the emitted photons due to the finite width of the slit.

In this section we quantify the variation of spectral line intensities as we observe the solar atmosphere in different positions and at different times. Also, we make estimates of what Lang et al.(1990) call the reproducibility of the observations and which we prefer to refer to as the *observational accuracy*, i.e. that part of the analysis which produces errors purely due to the factors involved in determining spectral line intensities by using a spectrometer on a spacecraft.

Line Position (Å)	log T_e(K)	%	Line Position (Å)	log T_e(K)	%
HeI 515.617	4.1	12	HeI 522.213	4.1	9
HeI 537.030	4.1	10	HeI 584.334	4.1	7
HeII 303.78^2	4.5	18	CIII 538.2^m	4.9	29
OIII 525.797	4.9	20	OIII 599.598	4.9	11
OIV 553.329	5.3	7	OIV 554.076	5.3	7
OIV 554.513	5.3	4	OIV 555.263	5.3	4
OIV 608.397	5.3	7	NeIV 543.891	5.3	18
NeIV 541.127	5.3	42	NeIV 542.073	5.3	7
OV 629.732	5.4	17	NeV 569.830	5.5	18
NeV 572.106^b	5.5	13	ArVII 585.75	5.7	24
NeVI 558.595	5.7	6	NeVI 562.7	5.7	10
NeVII 559.947	5.8	33	NeVII 561.378	5.8	17
NeVII 564.529	5.8	24	MgVIII 313.754	6.0	17
MgVIII 315.039	6.0	16	MgVIII 317.039	6.0	15
MgVIII 335.253	6.0	11	MgVIII 339.006	6.0	16
SiVIII 314.327	6.1	13	SiVIII 316.205	6.1	10
SiVIII 319.826	6.1	12	MgIX 368.057	6.1	12
AlX 332.789	6.1	23	AlXI 550.031	6.1	41
SiIX 341.949^b	6.2	2	FeX 345.102^b	6.2	9
FeX 345.723	6.2	10	FeX 365.543	6.2	10
CaX 557.74	6.2	12	CaX 574.01	6.2	13
MgX 609.76	6.2	13	MgX 624.94	6.2	6
SiIX 349.8^m	6.2	7	FeXI 341.113	6.3	12
FeXI 352.662	6.3	9	FeXI 369.153	6.3	13
SiX 347.403	6.3	14	SiX 356.0^m	6.3	14
FeXII 346.852	6.4	16	FeXII 352.106	6.4	13
FeXII 364.467	6.4	10	FeXI 358.621^b	6.3	39
SiXI 580.92	6.4	8	SiXI 303.78^2	6.4	34
SiXII 520.66	6.4	39	FeXIII 348.183	6.5	10
FeXIII 359.638	6.5	19			

Table 5.7: Lines selected for the study of statistical variance along with their temperature of formation and an estimate of the observational accuracy. Superscripts; m - multiplet, b - blend, 2 - second order line.

Table.5.7 presents the chosen line list and also shows the expected temperature of formation. This is taken as the peak of the ionisation balance curve for the emitting ion (Summers(1974)). Also shown in the table is a final estimate of the observational accuracy for each of the chosen lines (see below). The datasets used were the segments split and averaged as described in sec.5.3.1 and given in table.5.4.

For each line we plotted the log of the relative intensity against position number as given in table.5.4. That is, we normalised the uncalibrated intensities in counts, at each position number, to the average intensity of all position numbers. Note that the position numbers correspond to different summed sets rasters and different slit portions. Also, the different datasets were recorded on separate dates and varying times. The plots are presented in figs.5.20-5.22 and are ordered according to the temperature of formation of the line. For ions of the same element (assumed formed at the same temperature) they have been ordered by increasing wavelength. The line identifications and the window of the detector that recorded the data are given in the title of the individual plots. The plots are scaled appropriately for each line and so the y axis ranges are not all the same. This was done to allow any patterns to be more clearly visible. Also, error bars (output from the line profile fitting routine) are over plotted on each graph. In some cases the fits are so good that the error bars are difficult to see. These plots are similar to those presented by Lang et al.(1990) in their analysis of the CHASE measurements. Note that, as in table.5.7, some lines have superscripts denoting that they are multiplets or second order lines. Four lines are marked as blended. Their components are identified in sec.5.6.1 table.5.9. Each of these lines are included since the blended line's ions are formed at about the same temperature and so should show a similar pattern.

It is clear that a pattern of variation exists between lines formed at roughly the same temperature. This is not to say that the pattern reflects the temperature regime (that point will be discussed in sec.5.5.2) but simply that it may be possible to extend the CHASE analysis to CDS and so determine the statistical variation of the measurements. From fig.5.20-5.22 we can see a gradual development of the pattern from HeI 515.6$\AA$ through to NeVII 564.5$\AA$. This change is influenced mainly by the development of a sharp rise in intensity at position number 2 that remains high

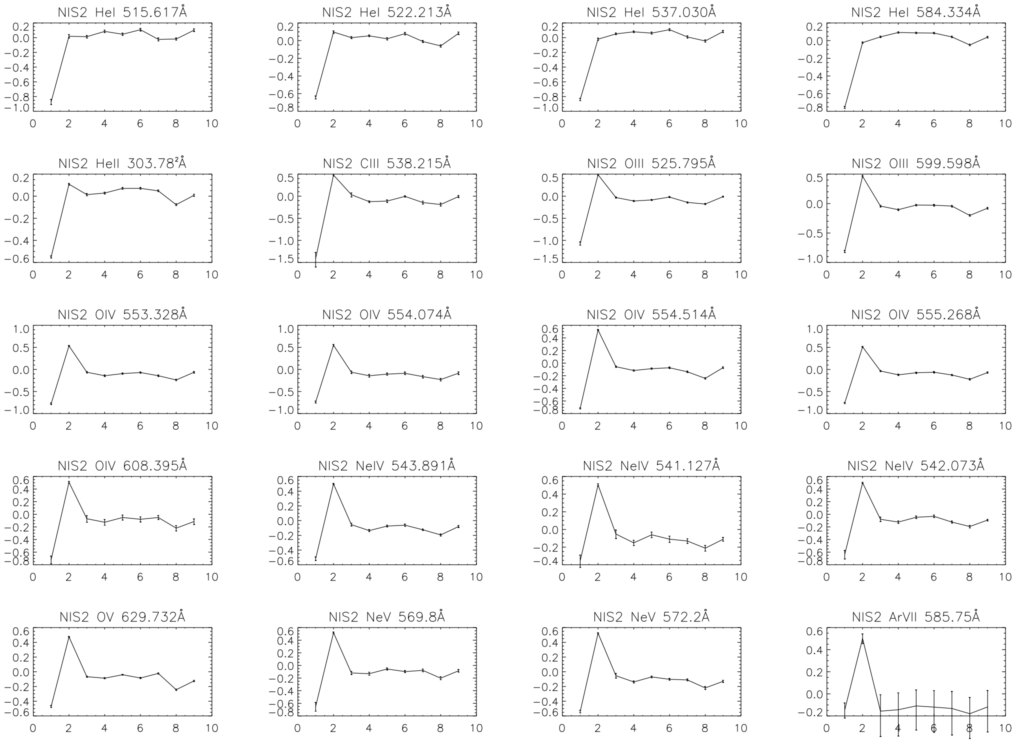


Figure 5.20: Plots of the logarithm of the relative intensity (see text) against position number (as defined in table.5.4). The line identification and the normal incidence extraction window that recorded it are noted above each individual graph. The plots are ordered by the temperature of formation of the line as given in table.5.4. Error bars are overplotted on each point and are largest for weak and blended lines.

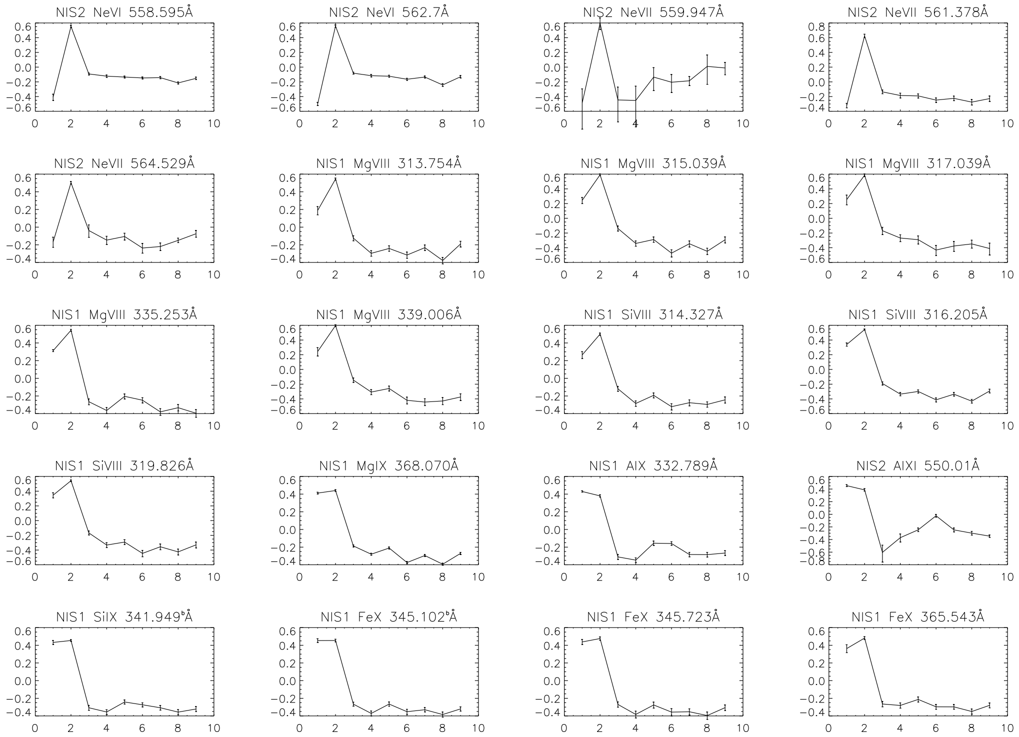

NIS2 NeVI 558.595Å
NIS2 NeVI 562.7Å
NIS2 NeVII 559.947Å
NIS2 NeVII 561.378Å
NIS2 NeVII 564.529Å
NIS1 MgVIII 313.754Å
NIS1 MgVIII 315.039Å
NIS1 MgVIII 317.039Å
NIS1 MgVIII 335.253Å
NIS1 MgVIII 339.006Å
NIS1 SiVIII 314.327Å
NIS1 SiVIII 316.205Å
NIS1 SiVIII 319.826Å
NIS1 MgIX 368.070Å
NIS1 AlX 332.789Å
NIS2 AlXI 550.01Å
NIS1 SiIX 341.949*Å
NIS1 FeX 345.102*Å
NIS1 FeX 345.723Å
NIS1 FeX 365.543Å

Figure 5.21: Plots of the logarithm of the relative intensity (see text) against position number (as defined in table.5.4). The line identification and the normal incidence extraction window that recorded it are noted above each individual graph. The plots are ordered by the temperature of formation of the line as given in table.5.4. Error bars are overplotted on each point and are largest for weak and blended lines.

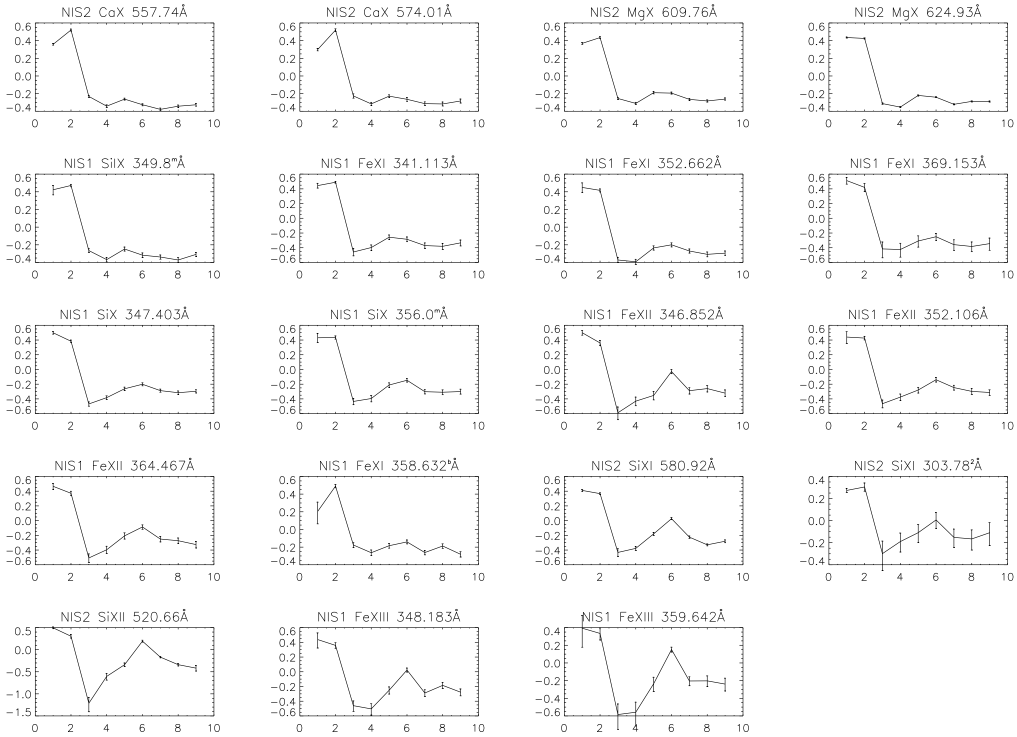


Figure 5.22: Plots of the logarithm of the relative intensity (see text) against position number (as defined in table.5.4). The line identification and the normal incidence extraction window that recorded it are noted above each individual graph. The plots are ordered by the temperature of formation of the line as given in table.5.4. Error bars are overplotted on each point and are largest for weak and blended lines.

through all the remaining plots. Position numbers 3 through 9 fluctuate around the mean up to HeII 303.8$\AA$ and then begin to be depressed to a slightly lower intensity. This is principally due to the influence, through the mean value, of the rise at position 2. Position number 1 varies more widely and in the plot of CIII 538.2$\AA$ is very much less intense despite the influence of position 2. By the time we reach the temperatures of the NeVII plots it has however begun to rise in intensity so that the sudden jump at MgVIII 313.7$\AA$ is not unexpected. From MgVII 313.7$\AA$ onwards the pattern has changed dramatically from HeI 515.6$\AA$ and position number 5 has also become more active. From here on there is a gradual development of the intensity of position number 6 until it rivals 5 in the AlX 332.7$\AA$ plot and briefly overcomes it in AlXI 550.0$\AA$. This latter pattern appears in greater evidence from FeXI 352.6$\AA$ onwards and raises the possibility that the temperature of formation of AlXI was wrongly assigned. In fact, results for AlX and AlXI were not presented by Summers(1974) and so results from ADAS were taken. From experience at looking at formation temperatures it is likely that had all the results been taken from ADAS the plots would have remained in order (except AlX and AlXI), but would have been shifted to slightly lower temperatures. From the plot of MgIX 368$\AA$ on, the intensity of position number 2 begins to drop slightly so that position 1 competes with it and dominates by the final four plots.

From the plots we see that the intensity of individual lines varies substantially from location to location and time to time with a typical variation being about a factor ten. However, variations of up to a factor fifty cannot be excluded, see for example, CIII 538.2$\AA$ or OIII 525.8$\AA$. In an attempt to be more quantitative and statistically accurate, we calculated the individual line's coefficients of count variation about their mean. We then then plotted these against the line's temperature of formation. This is simply the relative fluctuation of the counts about their mean, for a particular line, expressed as a percentage i.e.

$$\frac{\sqrt{\sum_i (F_i - \bar{F})^2/N}}{\bar{F}} \times 100.0 \tag{5.5}$$

where F_i is the fitted flux for the line in the ith dataset, N is the number of datasets,

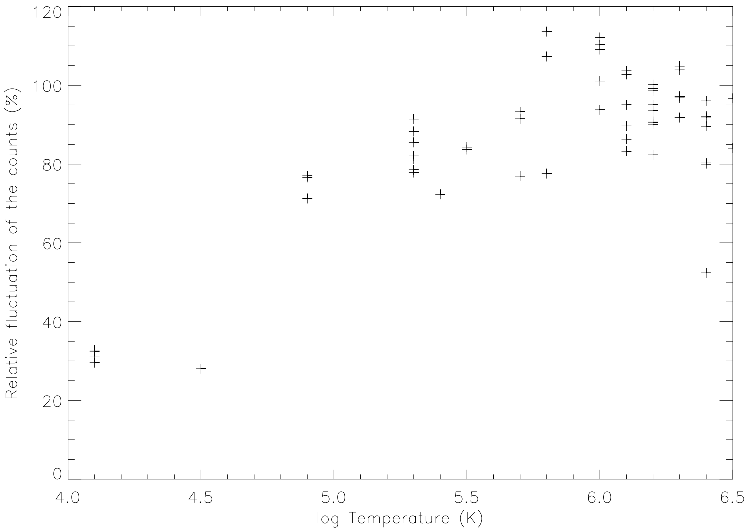


Figure 5.23: Relative fluctuation of the counts against temperature of line formation.

and the top line represents the standard deviation. Fig.5.23 presents the results. Statistically the meaning of this is that each point on the graph represents the percentage variability of the counts about their mean for a particular line. Therefore, we can see that the counts for individual lines can vary between $\sim 30\%$ and $\sim 110\%$ from their individual mean values.

Another way of expressing this variation is to take the standard error of the mean. For example, in the case of HeI 515.6$\mathring{A}$ the coefficient of variation is 32.8%, whilst the mean is 12496.8. Thus, the relative fluctuation is 0.328 and the standard deviation is $0.328\times12496.8= 4098.9$, giving a standard error of the mean ~ 1366.3 (i.e.$\sigma/n^{1/2}$). Therefore, we can say that if we have a normal distribution of intensities, the mean value of intensity for HeI 515.6$\mathring{A}$ is $1.25\pm0.14\times10^4$ and that if we took a larger number of measurements the mean would lie within this range (i.e. between $\sim 1.11\times10^4$ and $\sim 1.39\times10^4$) with a probability of 68% (Topping (1962)). This single result is presented more as an example to help with our understanding rather than a quantitative statement about the intensity variation of HeI in the solar atmosphere. However, it might be argued that the distribution of intensities is in some sense normal

Pattern	No. of Lines	Ions Included
1	5	HeI, HeII
2	11	CIII, OIII, OIV, NeIV
3	9	OV, NeV, ArVII, NeVI, NeVII
4	11	MgVIII, SiVIII, MgIX, AlX, AlXI
5	14	SiIX, FeX, CaX, MgX, SiIX, FeXI, SiX
6	9	FeXII, FeXI, SiXI, SiXII, FeXIII

Table 5.8: Ions contributing to each averaged pattern.

or averaged in the quiet sun, even including *overactive* and *underactive* regions as these might cancel each other to some degree. In any case, it is often usual to make this assumption at least about the variance of intensity, see for example, sec.5.2.1.

Given that the patterns appear similar for lines formed at the same temperature we decided to separate groups of patterns, as was done for CHASE, and form average patterns to try to determine how much each individual pattern of the group varies from the average. We believe this variation gives us an indication of the observational accuracy of the results. We argue that since the lines of the groups are formed at the same temperature and they are responding in a similar way to the variations in the emitting plasma, variations from the average at each point are indicative of variations due to the observational technique. In reality, we expect that solar conditions and the atomic model still play some part in determining these variations and so we consider the results of this analysis to be upper limits to the observational accuracy. The ions whose lines were selected for each group are given in table.5.8. To form the averaged plots we took the members of the six groups and found the mean of the intensities from each member at each position. This gave us a mean value for each position number from which we then found the mean of these means. Finally, we divided the means at each position by this last mean and plotted the logs of these values. In this way we produced six averaged patterns which are shown in fig.5.24. Fig.5.24 also shows the temperature spread of these averaged patterns. The features described in detail above are more evident in these plots. The sharp rise at position 2 appears quite convincingly in the second and third plots with the jump at position 1 and the

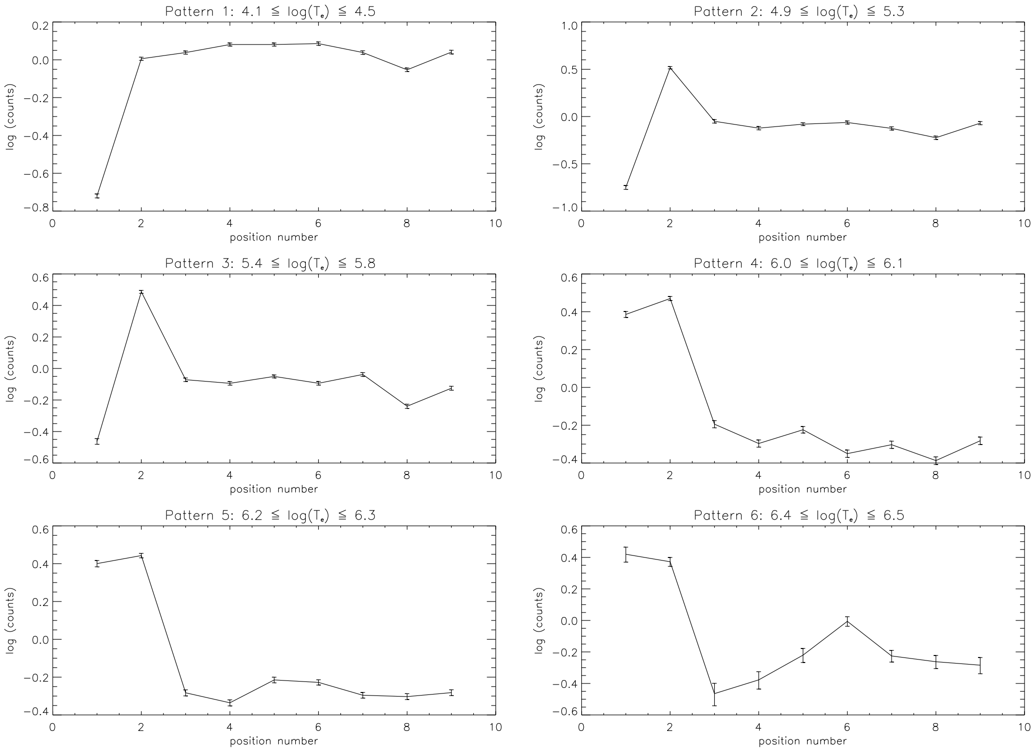


Figure 5.24: Averaged patterns corresponding to the groups of lines chosen (see table.5.8).

rise at position 6 evident in plots four and six respectively.

To calculate the observational accuracy we took the relative intensity values, used to produce the plots in fig.5.20-5.22, and calculated a mean from the values of the contributing patterns. This was done for each position. We then subtracted this mean from the relative intensities from the contributing patterns, at each position, and divided these by the mean. We then took the root mean square of these last values to give us our final results. The observational accuracy values obtained are shown in table.5.7 in the form of percentages. In addition, fig.5.25 shows the values plotted against line formation temperature. We can see that this measure of the observational accuracy varies between about 2% and 42% with the majority of values less than 20% and clustered around 10-15%. As mentioned these values are overestimates since it

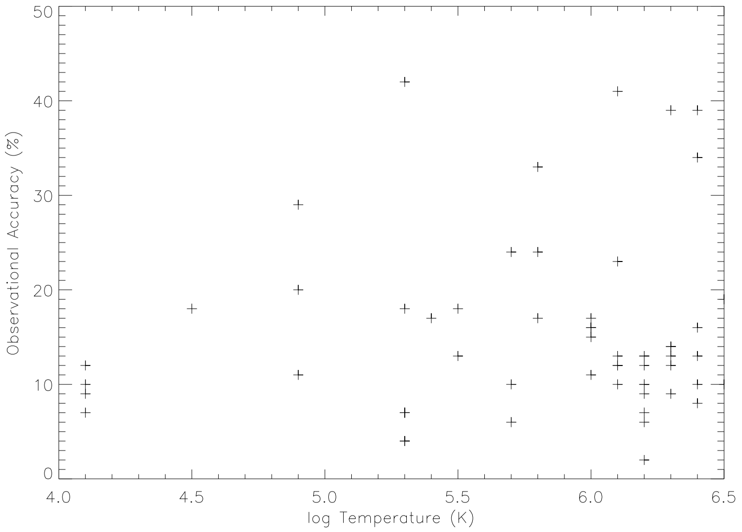


Figure 5.25: Observational accuracy of each line plotted against the log of the temperature of formation.

is difficult to completely exclude other influences (sec.5.5.1). Therefore, we conclude that the observational accuracy of the majority of these observations is in the region of 10-15% and we are cautiously optimistic that we may be able to work at around the 10% level or even better.

These results are similar to, or perhaps slightly better than, the conclusions of the CHASE analysis. In addition, it appears that the spectral line intensity variability is about the same as for CHASE. The determination of this observational accuracy is the main reason for looking at the raw data prior to any intensity calibration or atomic physics being applied. It would also be worthwhile to develop more accurate ways of determining the individual factors contributing to the observational accuracy.

Anomalies and Mis-Identifications

A number of important points arose during the above investigation. The original line identification list from the spectral atlas contained a further three lines which have been excluded from this study. These were identified as NeV 568.418$\AA$, MgVI 349.1$\AA$

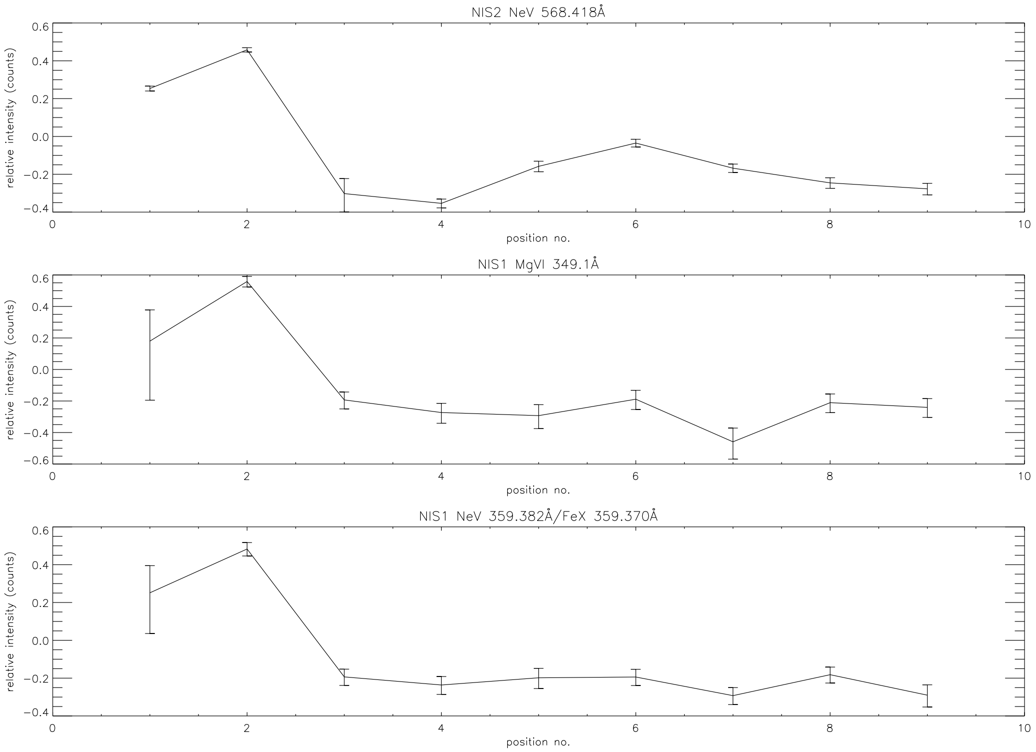


Figure 5.26: Three patterns which appeared odd for the reasons discussed in the text.

and a blend between NeV359.382$\mathring{A}$ and FeX 359.370$\mathring{A}$. The patterns of variation through the nine position numbers are given in fig.5.26. According to the temperatures of formation of NeV and MgVI (5.5 and 5.7 in the log respectively) the plots should appear between NeV 572.2$\mathring{A}$ and ArVII 585.75$\mathring{A}$ in fig.5.20. However, the high values of position number 1 suggest that the MgVI line should be associated with plots at a higher temperature of formation, perhaps closer to MgVIII or SiVIII i.e. 6.0 or 6.1 in the log. In addition, the rise in the intensity of position 6 in the NeV plot suggests that this line is formed at temperatures that are higher still (perhaps closer to MgX i.e. 6.2 in the log). For this reason we believe that these lines have been mis-identified. This points out a potentially valuable other use for the types of plots that we have presented. They may be able to confirm identifications and provide clues to the identity of unknown lines by allowing their temperatures to be estimated. No suggestion for what these particular lines are is given here.

Note also that the blended line appears to be more closely related to the pattern of FeX than NeV. This raises the possibility that the FeX line is the dominant component. However, since this is the only line in the study that is a blend of two lines formed at radically different temperatures, it was decided to reject it.

5.5.2 On the Physical Significance of the Patterns and the existence of a Temperature Dependence of Variability

The main purpose of producing the plots was to attempt to discover the statistical variation in the measurements due to factors resulting from the sampling and analysis methods. We expected patterns, similar to those seen by Lang et al(1990), to exist and that this would aid us in our study by grouping together lines formed at similar temperatures. While we do not expect the patterns to be reproduced exactly for lines formed at the same temperature, since there are many solar and atomic influences on the intensities, we believe that their existence is indicative of them being determined by systematic factors. As pointed out by Lang et al(1990) this suggests that there is hope of interpreting the intensities if we can discover the proper physical processes that contribute to them.

The analysis of CHASE measurements by Lang et al(1990) show a number of interesting results. They identified three distinct types of pattern that appeared to change sharply at definite temperatures. They also found an increasing variability with temperature. The latter was seen both for relative intensities with position and also in deviations from the averaged patterns i.e. for the highest temperatures, despite the most obvious change, the averaged pattern was the least secure in terms of applicability to other lines. Given the much larger number of lines available in the present study it seems worthwhile to re-examine some of these issues.

On the question of a temperature dependence of the variability fig.5.23 should give an indication. However, this diagram is inconclusive due to the small number of lines observed in the $1.2 \times 10^4 K$ to $6.3 \times 10^4 K$ region. There is some indication of an increasing slope with temperature and if the five points in the lower left corner are accepted then this idea is strengthened. There does appear to be a wider spread in fluctuation at higher temperatures but, again, this situation could be reversed if more points were added around the low temperature high variability region. We conclude that without more points it becomes a subjective exercise. It would be interesting to look again at this question with a specifically targeted set of lines that span a more controlled range of temperatures. Looking at the averaged patterns of fig.5.24 and the observational accuracies of fig.5.25 leads to the same conclusion. The variability appears higher for patterns 2 to 6 than for pattern 1 and there is a greater scatter in the values of observational accuracy at higher temperatures. However, if the lines contributing to pattern 1 are not a representative sample of the low temperature variability then much more cannot be said. In any case, the activity of points 1 and 2 partly controls the variability for this data collection.

It is interesting to compare the temperatures at which the averaged patterns change in the present work with the results presented by Lang et al.(1990). In their study they found a pattern change at $\sim logT = 5.5$ to 5.8 and a more distinct one at $\sim logT = 6.1$. They suggest that the patterns are associated with the physical structure of the atmosphere at different temperatures. That is, the first shows grossly the supergranulation structure of the network in the chromosphere and the second shows the loop structures of the higher temperature lines in the transition zone and corona. They also suggest that the third pattern may be evidence for an additional transition in the solar atmosphere at this temperature. With regard to this last point, they suggest that the discontinuity at $logT = 6.1$ could provide supporting evidence for the type of cool loop model proposed by Antiochos & Noci (1986). Indeed, a discontinuity of this type may be supportive of the arguments of Feldman & Laming (1994).

In this work, we found the most distinctive pattern changes between 1 and 2 and then between 3 and 4 of fig.5.24. As mentioned, these appear influenced by the activity of points 2 and 1 respectively. In addition, they are not as 'sharp' as those seen by Lang et al.(1990). These alterations occur at temperatures between $\sim logT = 4.5$ to 4.9 in the first case and between $\sim logT = 5.8$ and 6.0 in the latter case. The absence of a distinctive change at $logT$=4.5 to 4.9, in the CHASE paper, and at $logT$=5.5 to 5.8, in the present work, suggests that there may be no deeper physical significance to the sharp pattern changes. However, the sharpest change occurred between $\sim logT = 5.8$ and 6.0 which is closer to that seen with CHASE. Also, if the changes correspond to particular features of the atmosphere it is possible that the temperatures at which the changes occur are dependent on the shapes and sizes of the features 'caught' by the spectrometer on the day. However, this is equivalent to saying that the pattern changes fall within the limits of the natural variation of the atmosphere. The observations with the grazing incidence spectrometer long slit should precisely match the observations done with CHASE and clear up these speculative queries. It is recommended that, for this purpose, the same source as used by Lang et al.(1990) is taken to decide the temperatures of line formation.

5.6 Demonstration of Method and Reconstruction of the DEM for CDS-NIS

As a demonstration and test of the analysis procedures we used the normal incidence spectrometer data from the DEMST_1 observations taken on September 22nd 1996. This demonstration provided an opportunity to re-assess each of the software components and to test their robustness. Modifications and adjustments were made as a result of this exercise that have subsequently been included in the routines described above. In addition, potential problems with atomic data and line blending were identified examination of which is now part of the final study.

5.6.1 Preparation of Intensity File

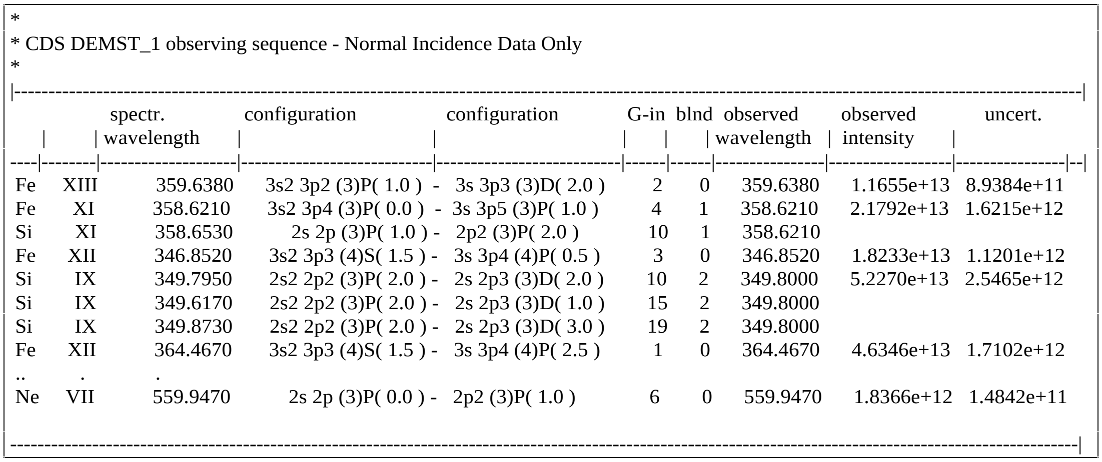
Each of the lines identified for the study of statistical variability in section.5.5.1 table.5.7 were considered for this preliminary DEM attempt. Fifty nine lines had been identified including three multiplets and four blends (see section.5.5.1). Using the preliminary spectral atlases for NIS1 & NIS2, completed as part of the initial study of the data (see section. 5.3.1), we identified those components of the three

Main Component	Actual/Other Contributors
SiIX 349.8^{m}Å	SiIX 349.795Å SiIX 349.617Å SiIX 349.873Å
SiX 356.0^{m}Å	SiX 349.05Å SiX 349.01Å
CIII 538.2^{m}Å	CIII 538.080Å CIII 538.149Å CIII 538.312Å
SiIX 341.949^{b}Å	MgVIII 342.062Å
FeX 345.102^{b}Å	SiIX 344.951Å SiIX 345.124Å
FeXI 358.621^{b}Å	SiXI 358.653Å
NeV 572.106^{b}Å	NeV 572.336Å

Table 5.9: Line components contributing to the make up of the blended lines and multiplets from table.5.7. Superscript m - multiplet, b - blend.

multiplets that contributed to the formation of the line shapes that were actually fitted. In addition, the atlases were used to identify those lines that were expected to contribute as components of the four blends. Table.5.9 lists the seven blends and multiplets and the components that were identified in each case. Due to the possibility of opacity affecting transitions within ions of moderate charge (Doyle & McWhirter(1980)), we rejected the five helium lines from the analysis. In addition, a second order intensity calibration was not available at the time of the analysis so the SiXI 303.78Åline was excluded. This left a total of fifty three lines, not including the blend/multiplet components.

The *s4878r03* data of the DEMST_1 observations was then processed, corrected and fitted as described in section.5.4.2. The routine *i2dem.pro* was used to create the intensity file, a segment of which is shown in fig.5.27. As mentioned in section.5.4.2, the ion identifiers, observed and spectroscopic wavelengths, observed intensities and uncertainties are printed directly by the routine. The *configurations* are not essential for further processing and are entered by hand. A blend index, *blnd*, is also entered



```
*
* CDS DEMST_1 observing sequence - Normal Incidence Data Only
*
|--------------------------------------------------------------------------------------------------------------------------------------------|
           spectr.        configuration              configuration             G-in  blnd  observed      observed          uncert.
    |      | wavelength   |                          |                         |     |     | wavelength  | intensity       |
----|------|--------------|--------------------------|-------------------------|-----|-----|-------------|-----------------|---------------|--|
 Fe   XIII    359.6380    3s2 3p2 (3)P( 1.0 )  -  3s 3p3 (3)D( 2.0 )     2     0    359.6380    1.1655e+13  8.9384e+11
 Fe     XI    358.6210    3s2 3p4 (3)P( 0.0 )  -  3s 3p5 (3)P( 1.0 )     4     1    358.6210    2.1792e+13  1.6215e+12
 Si     XI    358.6530       2s 2p (3)P( 1.0 ) -  2p2 (3)P( 2.0 )       10     1    358.6210
 Fe    XII    346.8520    3s2 3p3 (4)S( 1.5 ) -  3s 3p4 (4)P( 0.5 )      3     0    346.8520    1.8233e+13  1.1201e+12
 Si     IX    349.7950    2s2 2p2 (3)P( 2.0 ) -  2s 2p3 (3)D( 2.0 )     10     2    349.8000    5.2270e+13  2.5465e+12
 Si     IX    349.6170    2s2 2p2 (3)P( 2.0 ) -  2s 2p3 (3)D( 1.0 )     15     2    349.8000
 Si     IX    349.8730    2s2 2p2 (3)P( 2.0 ) -  2s 2p3 (3)D( 3.0 )     19     2    349.8000
 Fe    XII    364.4670    3s2 3p3 (4)S( 1.5 ) -  3s 3p4 (4)P( 2.5 )      1     0    364.4670    4.6346e+13  1.7102e+12
 ..      .    .
 Ne    VII    559.9470       2s 2p (3)P( 0.0 ) -  2p2 (3)P( 1.0 )        6     0    559.9470    1.8366e+12  1.4842e+11

--------------------------------------------------------------------------------------------------------------------------------------------|
```

Figure 5.27: Example of DEM intensity file format using DEMST_1 data.

by hand. Zeros are placed against every line that is to be treated separately, while blended lines are grouped together using the same index. The observed intensity for the blended lines is specified next to the main component and the contributing components' intensities are left blank (e.g. SiXI in figure). When the grouping of blended lines takes place the blanks are taken as zeros and the total remains as that of the main component. The same method is used for members of multiplets. Finally, an index number, *G-in*, referencing the appropriate G(T_e) function in the 'kernel collection file' must be entered.

5.6.2 Preparation of Kernel Function Collection File

G(T_e) functions must be available for each line in the intensity file. The G(T_e)'s are stored in 'collection files' which are output from ADAS routine ADAS506 (see Summers et al.(1996)). Each G(T_e) in the collection file has a specific transition number and this is used to provide the connection with the observed intensity (the *G-in* index mentioned above). Collection files are built up by selecting the appropriate transitions from G(T_e) functions stored in the ADAS database. G(T_e) functions are stored for a wide range of ions at uniform electron densities and pressures. For this demonstration it was necessary to locate G(T_e) functions for the radiative transitions

Ion
NeIV, ArVII, NeVI(558.595Å only), AlX, AlXI, FeX, CaX, FeXI, FeXII, FeXII

Table 5.10: List of ions where data was not available in ADAS and $G(T_e)$ functions had to be generated specifically. The data used were from lower quality Born calculations. Exceptions were FeXII, NeVI & FeXIII where high precision calculations of Tayal, Henry & Pradhan (1987), Zhang, Graziana & Pradhan (1994) and Fawcett & Mason (1989) became available.

forming each line in tables.5.7 and 5.9.

Given the problems associated with deriving plasma structure and the necessity for minimising errors in the atomic data, we wish to use only the highest quality data for our complete DEM study. However, for this demonstration we relax these constraints and ensure that we have adequate coverage for all the lines in the intensity file. This includes high precision electron collisional excitation data but also lower quality 'fill in' data generated by simpler methods. Most of the data were taken from ADAS where a large database of $G(T_e)$ functions (*adf20* files) have already been generated over a large range of constant densities ($1 \times 10^8 \rightarrow 1 \times 10^{12} cm{-}3$) and constant pressures ($1 \times 10^{10} \rightarrow 1 \times 10^{16} Kcm^{-3}$). The electron collisional excitation and spontaneous radiative data sources are given in the relevant files and also by Summers(1996). For some transitions data is sparse and certainly not of the highest quality. Where necesssary we generated $G(T_e)$ functions specifically for this task based on Born approximation calculations carried out at JET. These data do not appear in the ADAS central database and the relevant ions are listed in table.5.10. The ionisation balance calculations of Arnaud & Rothenflug (1985) with a density dependent correction based on the collisional-radiative models were used in the generation of the G(T_e) functions.

Using ADAS506 a kernel 'collection file' was created for a uniform pressure of $3 \times 10^{15} Kcm^{-3}$. This is shown in fig.5.28. The temperature coverage extends from a few tens of thousands to tens of millions of degrees kelvin. This is an adequate range for solar atmospheric studies, encompassing the upper chromosphere through

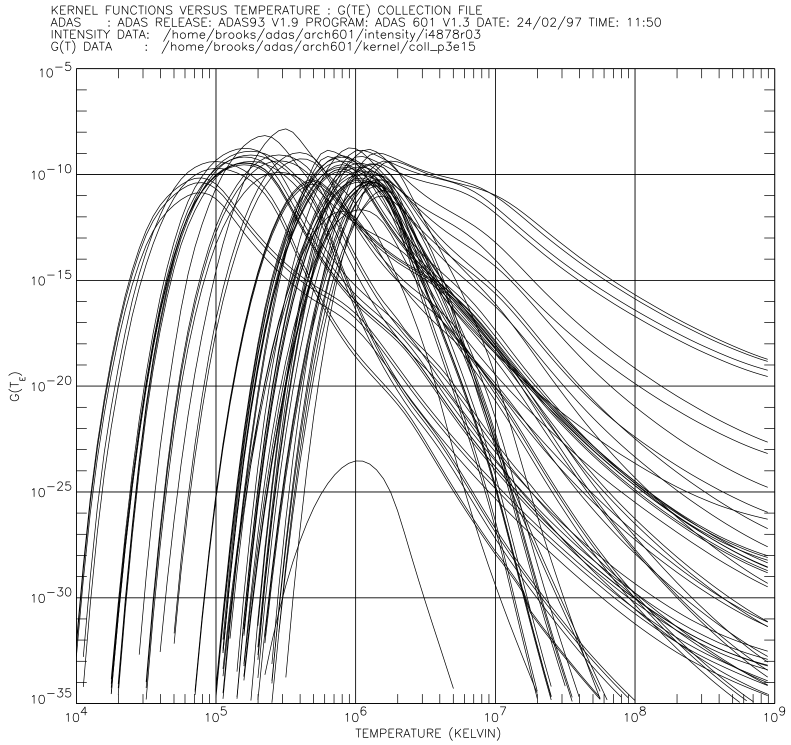


Figure 5.28: Kernel function (G(Te)) collection file for the DEM demonstration run.

the transition region to the corona. Note that there is particularly dense coverage around 10^6K. These lines come principally from the NIS1 extraction window and are thus more easily identifiable as a cluster in the DEM (see below). This allows us to make a more immediate judgement about the range of values associated with each window.

5.6.3 Sample Reconstruction of the DEM

To reconstruct the DEM distribution we require observed intensity, predicted intensity and elemental abundance information. These are input to specific ADAS format files and then used in conjunction with ADAS601 to perform the integral inversion. The prepared intensity and kernel collection files, described above, were used along with the photospheric abundances of Grevesse & Anders (1991). The values of these elemental abundances are shown in table.5.11.

Numerical Issues

Solutions to the bi-variate intensity integral equation (eq.2.61) are rare in astrophysics. However, many techniques have been developed and thoroughly tested to solve eq.2.62. ADAS601 uses the algorithm of Thompson(91) and the method is outlined in detail by Harrison & Thompson (1992) and by Summers et al.(1996). Here we give a brief overview.

Thompson used a *data adaptive smoothing approach* which required the equation 2.62 to be discretised as,

$$I_{i\to j} = \frac{1}{4\pi}\frac{N_{tot}}{N_H}\sum_{k=1}^{N}\phi_k^* \int_{t_{k-1}}^{t_k} G_{i\to j}(T_e)\phi_0(T_e)dT_e \tag{5.6}$$

where the transition is from the upper level i to the lower level j and we have ignored the dependence of the G functions on density. The G function is the kernel in this case. $\phi_0(T_e)$ is an initial estimate of of the expected gross behaviour of $\phi(T_e)$ which reduces the dynamic range of the problem before the full solution is obtained. $\phi_0(T_e)$ is named the 'prior' function. Also, ϕ_k^* represents a weighted average of the ratio of the

Element	Symbol	Abundance
Hydrogen	H	1.0
Helium	He	9.8×10^{-2}
Carbon	C	3.6×10^{-4}
Nitrogen	N	1.1×10^{-4}
Oxygen	O	8.5×10^{-4}
Neon	Ne	1.2×10^{-4}
Sodium	Na	2.1×10^{-6}
Magnesium	Mg	3.8×10^{-5}
Aluminium	Al	3.0×10^{-6}
Silicon	Si	3.6×10^{-5}
Sulphur	S	1.6×10^{-5}
Argon	Ar	3.6×10^{-6}
Calcium	Ca	2.3×10^{-6}
Iron	Fe	4.7×10^{-5}
Nickel	Ni	1.8×10^{-6}

Table 5.11: Table of photospheric abundances used in the DEM demonstration. The data are from Grevesse & Anders(1991) and only include elements whose abundance is greater than 10^{-6} of hydrogen. It is possible that some elements may have enhanced abundances in the corona (see discussion sec.5.6.5).

full solution $\phi(T_e)$ to the prior function, $\phi_0(T_e)$, between the temperature extremes of the components of the numerical integral i.e. t_{k-1} and t_k. Thus,

$$\phi_k^* \equiv \frac{\int_{t_{k-1}}^{t_k} G_{i\to j}(T_e)\phi(T_e)dT_e}{\int_{t_{k-1}}^{t_k} G_{i\to j}(T_e)\phi_0(T_e)dT_e} \tag{5.7}$$

The method effectively scales the solution using the prior function and has proven to be successful in preventing over-smoothing at high and under-smoothing at low values of the DEM.

The prior function is evaluated by approximating the G function as the product of its integral with the Dirac delta function of the difference between the temperature and the temperature of the maximum of the G function viz.

$$\begin{aligned} G_{i\to j}(T_e) &= \bar{G}_{i\to j}\delta(T_e - T_{i\to j}^{max}) \\ &\equiv \int_0^\infty G_{i\to j}(T_e)dT_e\delta(T_e - T_{i\to j}^{max}) \\ \bar{\phi}(T_{i\to j}^{max}) &= \frac{4\pi N_H}{N_{tot}}\frac{I_{i\to j}}{\bar{G}_{i\to j}} \end{aligned} \tag{5.8}$$

The values $\log(\bar{\phi}(T_{i\to j}^{max}))$ are the first approximations to the solution and a smooth representation of the prior function is obtained by fitting a cubic B-spline to them. This method is satisfactory only between the minimum and maximum values of $T_{i\to j}^{max}$ so a linear extrapolation is made to lower and higher temperatures. The methods used and further detail of the above description are given by Summers et al.(1996).

First Attempt

The sample DEM is shown in fig.5.29. The crosses are the first approximations calculated for each of the observed lines. The dashed line through these points is the prior function and the solid line is the final DEM. Note that the temperature maxima of the G(T_e) functions will not correspond exactly to the values quoted in table.5.7 for each line. Thus the temperature coverage extends from ~ 5.0 → 6.3 in the logarithm only. Note also that the solution outside this temperature range has no physical meaning.

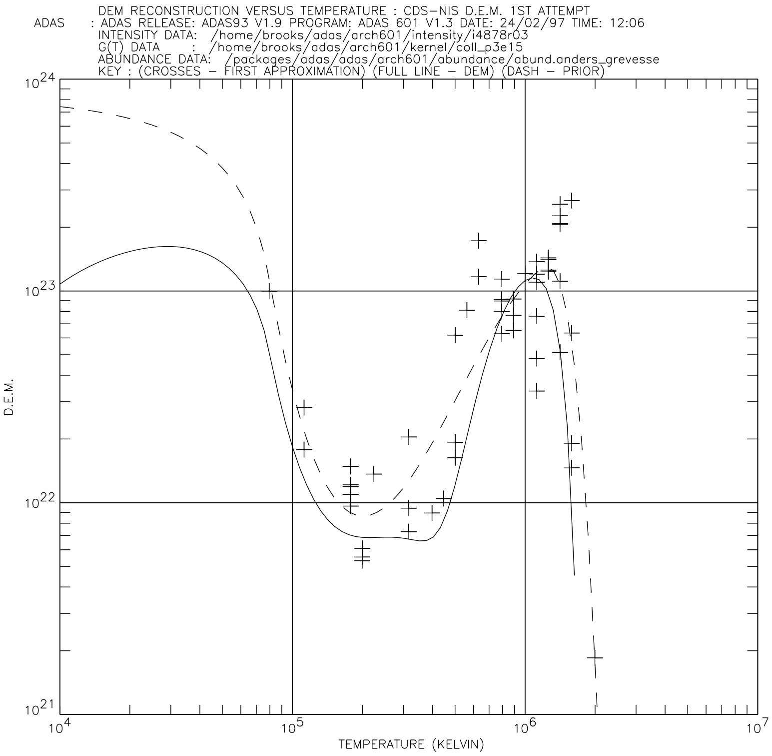


Figure 5.29: The emission measure differential in temperature computed from CDS-NIS data for a uniform pressure of $3 \times 10^{15} K cm^{-3}$.

Within the range, the overall shape of the curve is largely familiar with the exception of the unphysical kink around $3-4\times10^5$K. Variations of this kind are unrealistic as we cannot reasonably expect to resolve structures that vary more sharply in temperature than the width of the G(T_e) functions used. The slope down from around 10^5K to form the shallow dip between 10^5 and 10^6K, followed by the peak around 10^6K is largely consistent with the results of previous authors although the actual numerical values of the DEM vary (see for example Brosius, Davila, Thomas & Monsignori-Fossi(1996)). The results of Lang et al.(1990) were consistent with this also except that they found a second 'hump' around 3×10^5K. However, as suggested by Thompson (1991) this 'hump' was overly influenced by the OVI($2s\ ^2S_{1/2}-2p\ ^2P_{3/2}$) 1031.91Å line and he was able to remove the behaviour by giving this line less weight in the inversion. This is indicative of a significant error in the $G(T_e)$ function for that line. Errors in the G(T_e) functions may also be responsible for the kink in our curve. Alternatively, it may be due to the fact that the atomic data used, for some of the lines, was not of the highest quality.

We attempted to identify the line or lines responsible for this behaviour. By a process of line elimination and DEM re-evaluation we came to the conclusion that a smooth distribution, within the temperature range, could be recovered by excluding the two NeVI lines from the analysis. This might indicate again that there is some problem with the atomic data for this ion. It is essential to make a realistic assessment of the errors before attempting a proper DEM. In this demonstration we used the estimated standard errors output from the line profile fitting procedure. A more meaningful estimate would be to take the larger percentage between this value and the value of the observational accuracy for the line determined from the methods of sec.5.5.1.

5.6.4 Sensitivity of Spectral Lines Recorded by CDS-NIS

Thompson (1991) suggested that one way to ensure the G(T_e) functions do not contain significant errors would be to remove individual lines from the dataset and compute a DEM distribution from which the line intensity can be predicted. A comparison

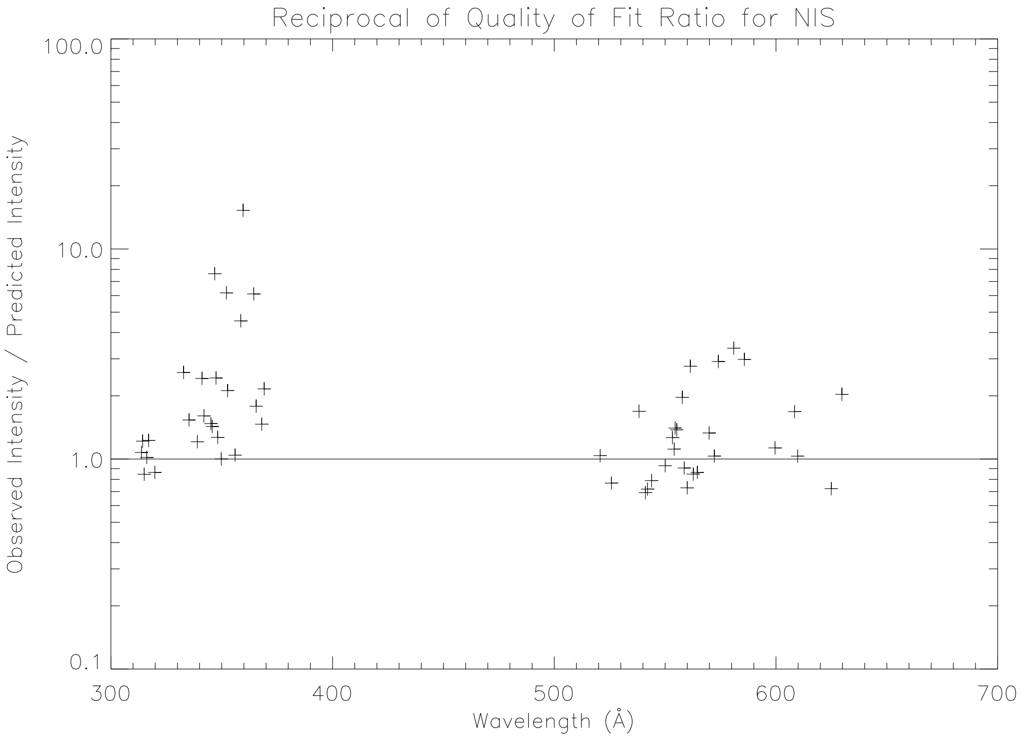


Figure 5.30: Ratio of observed to predicted intensities for NIS. NIS1 is the shorter of the two wavelength regions.

between the predicted and observed intensity would then show up any discrepancy which would identify any potential errors in the G(T_e) functions. Lang et al. (1990) presented something similar which they called quality of fit ratios, *qfr*. These are ratios of intensities predicted by the DEM code to the observed intensities and are calculated for all lines in the dataset without omission. These *qfr's* were used extensively by Lang et al.(1990) to try to obtain information on, for example, the correctness of the elemental abundances, errors introduced by the size of the recorded signal, errors related to the photometric calibration, etc.

We are also concerned with the possibility of a revision of the pre-flight intensity calibration (see sec.5.4.2). At time of writing there was some suggestion (Lang(1997) - private communication) that the NIS1 band was over sensitive. A glance at fig.5.29

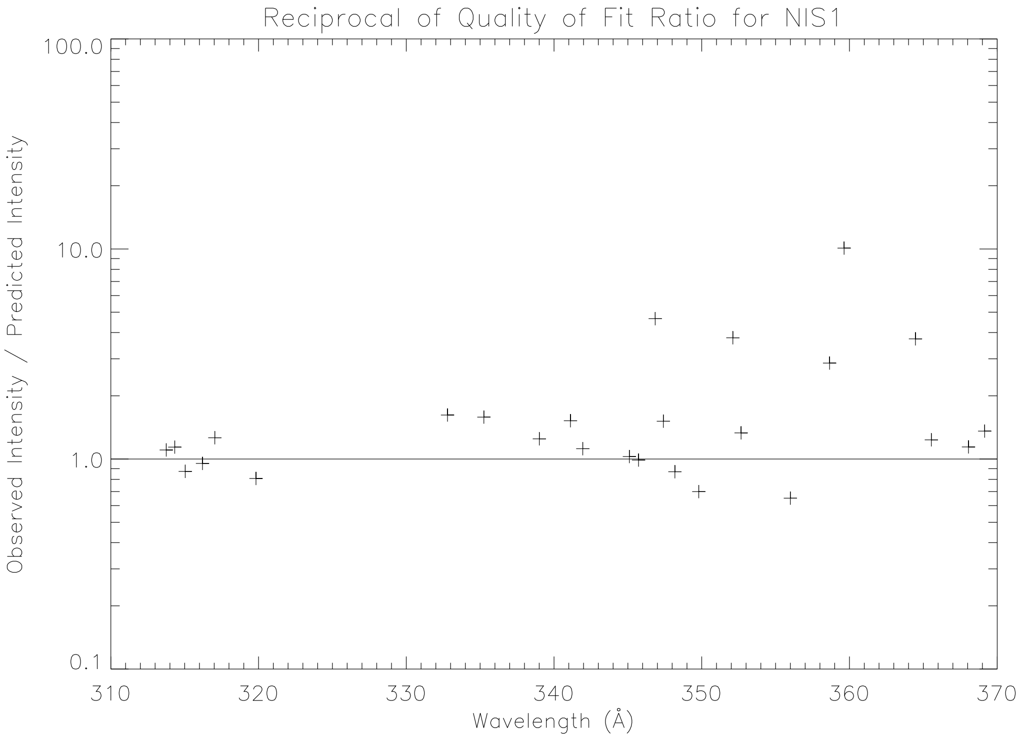


Figure 5.31: Ratio of observed to predicted intensities for NIS1.

shows a large cluster of values around 10^6K and $10^{23} cm^{-5} K^{-1}$ in the DEM. This temperature region corresponds mainly to the NIS1 data. Specifically, there is a large spread of values of the DEM for a group of lines formed at the same temperature (note the five points that cover the range $\sim 3 \times 10^{22} - 1.3 \times 10^{23} cm^{-5} K^{-1}$ and appear in a vertical line at $\sim 1.1 \times 10^6 K$).

To give clues to the accuracy of the calibration and the atomic data we plotted the reciprocal of the *qfr*'s against wavelength (see fig.5.30). The results for NIS2 are markedly better than for NIS1 and there does not appear to be any noticeable wavelength dependence of the *qfr*'s which might have helped in identifying any problems with the instrument sensitivity in particular spectral regions. Observed intensities from NIS1 appear too high by up to at least a factor ten while NIS2 is better showing an apparent overestimate of only up to a factor four. It is unlikely that this is a true reflection of the sensitivity of the detector since the outliers may be affected by poor fits to the data, inadequate treatment of weak and blended lines that may remove counts from the total fluxes and poorer quality atomic data. For the final study such outliers will be investigated individually. As an example, fig.5.31 shows the same graph but excluding contributions from NIS2. That is, the approximate DEM is computed self consistently using the values for NIS1 only. Thus any disparity between the calibration of the two extraction bands should be removed. A group of five prominent outliers remain whose intensities are still too high compared to those predicted but, encouragingly, the disparity is reduced from previously. These were identified and are given in table.5.12 along with the reciprocal of the *qfr*'s for both examples. It is interesting that the largest discrepancies occur only for iron lines. This raises the possibility that there is some error in the elemental abundance. However, after investigation other more likely explanations arose. For example, the FeXII 352.106Åwas found to be blended with the MgV($2p^6\ ^3P_1 - 2s2p^5\ ^3P_0$) 352.201Å line which obviously causes an overestimation of the flux during the fit. Back tracking through the analysis procedures pointed up that there had been difficulty in unambiguously separating the FeXIII 359.638Åline from two neighbouring lines (FeX 359.37Å& NeV 359.382). Finally, the G(T_e) function used for the FeXI component of the blended line was generated from Born approximation data and so is not of the highest quality.

It is clear that some re-working and scrutiny of problem lines in conjunction with checks on predicted intensities and atomic data will be essential for dealing with such cases in the final DEM study. In this way we hope to close the gap as closely as possible between the theoretical and observed intensities. If we exclude the five lines from table.5.12, the observed intensities are well within a factor two of the predicted ones which is encouraging for this approach.

5.6.5 Comment on Solar Atmospheric Structure Determination

Since we have been able to derive an initial, and crude, differential emission measure distribution it is appropriate here to comment further on the objectives of the

Identification	1/qfr(using NIS)	1/qfr(using NIS1 only)
FeXIII 359.638Å	15.3	10.1
FeXII 346.852Å	7.6	4.7
FeXII 364.467Å	6.2	3.7
FeXII 352.106Å	6.2	3.8
Blend FeXI 358.621Å/ SiXI 358.653Å	4.5	2.9

Table 5.12: Ratios of observed to predicted intensities for the five outliers of NIS1. Results are quoted for the DEM computed with both NIS1 and NIS2 and also for NIS1 only.

complete study; what we hope to achieve and how to proceed. This follows from the discussion of chapter.2 section.2.2.3. With an accurate intensity calibration, best available atomic data and a realistic assessment of the errors we should ultimately be able to say definitively whether the DEM approach tells us some useful properties of the solar atmosphere. Clearly it does as much progress is likely to be made in areas such as line identification or radiative transfer computations by utilisation of the DEM technique. The expectation is however that disparities between observation and theory will remain independent of the quality and accuracy of the data and techniques applied. For example, Lang et al.(1990) found that the variability of their DEM 'solutions' that satisfied the integral inversion was greatest in the temperature regime where their confidence in the observational and atomic data was greatest.

Such a conclusion signals either the breakdown of assumptions underlying the theory of the DEM method, or mis-assessment of the errors in the observational and theoretical data. Lang et al.(1990) discussed the various experimental and theoretical factors that might cause one or other of the problems mentioned above. Subsequently these suggestions were revisited by Judge, Woods, Brekke & Rottman (1995). Here we will address some of these authors' points again.

In this chapter, and in the on-going study as a whole, we have tried to reduce the possibility that data errors (atomic and observational) have a significant effect on our analysis. The raw quality of the observational data was assessed in sec.5.5.1. Problems such as mis-identification of lines were partly eliminated through the CHASE pattern analysis (see e.g. sec.5.5.1). Effects of the inaccurate treatment of blended lines and background signal subtraction cannot be completely avoided but the development of the maximum likelihood line profile fitting routine at least quantifies the errors associated with each fit. Also, as suggested above, the scrutiny of individual lines using an approximate DEM can be useful in this respect. In general, with the exception of badly blended lines, these effects were less than the statistical values of the observational accuracy found by the methods of sec.5.5.1. The assessment of these observational accuracies also partly eliminates many of the queries raised by Lang et al. with regard to spacecraft transmission errors etc. The routine adopted for the inversion procedure has been shown to be robust both in the assessment exercise (Harrison & Thompson (1992)) and in the analysis of SERTS data (Lanzafame et al. (1997)). Also, the data analysis procedures have now been well tested as a result of this thesis demonstration.

Many of the techniques developed will be used in much greater detail for the full study but there are a number of further points in this area that were not completely addressed. Clearly, there is an anxiety associated with the lack of a definitive pre-flight intensity calibration (see 5.4.2 and refsec:chap-564). A document is in preparation (Kent et al.(1997)) and some work is in progress investigating quality of fit ratios for the pre-flight calibration lines. Also, fig.5.30 is not indicative of any gross disparity between NIS1 and NIS2. Investigation of *qfr*'s for lines with highest quality atomic data may alter this situation. It is also possible that a definitive statement will not be made before the relaunch of the SERTS rocket, scheduled for autumn 1997. The problem of providing atomic data of best quality was not fully addressed. ADAS contains the best available data but further work may need to be done, for example on S-like and Cl-like Fe, where observations were available but atomic data was lacking. However, ADAS is in a uniquely advanced state to provide back up for this. It is also possible that errors in the elemental abundances used may affect the final study. There is a widely held view (see e.g. Laming, Drake & Widing(1995)) that the abundances of elements with a low First Ionisation Potential (e.g. sodium, calcium) are overabundant in the corona by $\sim$ factor of 3. These authors suggest there may be a diffusion process which tends to remove ions from the photosphere to the corona.

Elements that are difficult to ionise at the lower temperature of the photosphere would be partly left behind in this scenario. If this is correct then the abundances used in our demonstration DEM will be incorrect. However, this suggestion is not fully accepted and it is possible that the observational errors are too large to draw any conclusions on this so called FIP effect. Clearly if the errors are so large as to potentially obscure a physical process then more detailed analysis is required.

The possibility that there is a failure in one or more of the assumptions underlying the DEM method can only be concluded after the full study is completed with minimal errors. Of the explanations raised by the above authors we can however make some suggestions.

In the above DEM demonstration we excluded the helium lines on the grounds that their intensities may be altered by opacity. This is a concern for many observed lines. A paper is in preparation, Summers et al.(1997), which extends the analysis of Doyle & McWhirter(1980) from lines which show opacity effects as we view them across the solar limb, to the case of lines that are optically thick on the solar disc. Thus a complete prediction of the optical thickness of lines can be made allowing us to exclude these from the DEM analysis. Clearly there are optical thickness effects that must be considered for the cooler denser chromospheric region which may restrict the applicability of the DEM approach.

In the transition zone-corona region the results presented in sec.3.2.1 and sec.5.3.1 are consistent with a failure of the assumption of equilibrium for the ionisation balance. Dynamic effects clearly require that metastable states are fully accounted for in the atomic calculations but this falls short of computing a full nonequilibrium solution such as that of Spadaro and co-workers. As mentioned in chap.4 sec.4.7 we have the capability to examine such issues using the dynamic ionisation codes of ADAS. However, it has not yet been fully evaluated how such a model can be incorporated into the DEM method.

Evidence for disequilibrium in emission measure analysis of Li-like and Na-like isoelectronic sequences was reported by Judge et al.(1995). They observed that lines of these sequences are systematically underestimated by theoretical calculations. This is indicative of dynamics because Li-like and Na-like ions suffer from long recombination times from He-like and Ne-like ions where larger than average temperature increases are required to open the closed $n = 1$ and $n = 2$ shells (see chap.3 sec.3.2.1). However, it is not clear whether this effect is significant as it was not seen by Lanzafame et al.(1997) in their analysis of SERTS data. A dynamical ionisation balance calculation is required but how large the disequilibrium is and whether it can be detected in the sort of 'averaged' structures of a DEM analysis is not yet known.

Invoking assumptions about surfaces of constant temperature and density in order to cast the integral in the form of eq. 2.62 is also questionable. Dynamic activity, flows, spicules etc. produce steep rapidly changing temperature gradients. In addition, Gabriel(1976) showed that there are strong magnetic fields around network features in the chromosphere which decrease in strength rapidly with height. The assumption that surfaces of constant plasma density and temperature both exist in the solar atmosphere and coincide is highly debatable.

Lang et al.(1990) and Judge et al.(1995) cite a failure of some aspect of the atomic model as another possibility. As mentioned by them, inaccurate treatments of dielectronic recombination and ionisation coefficients and, in particular, additional density dependences in them should be modelled more precisely. Judge et al.(1995) admit that they cannot completely discount an inadequate treatment of dielectronic recombination in their own work. One of the strengths of our work is in this area (see chap.4) so this problem has been and will continue to be addressed. Lanzafame et al.(1997) interestingly find density dependencies in lines where it would perhaps not be expected. The use of ADAS alleviates many of these difficulties as it provides a complete density dependent collisional-radiative model. Also, the SOHO-CDS & SUMER DEM study should provide new insight.

This leads us to another important point. Recent work by Judge et al.(1997) suggests it is debatable whether a solution to the formal mathematical extension of the DEM distribution to two independent variables, temperature and density, can be reconstructed from the observations at all. There is much to be investigated in this area. It remains to be seen whether a better inversion method than that of Hubeny & Judge(1995) would help in the derivation of the solution. As pointed out by Judge et al.(1997), some success has been achieved with inversion problems by tailoring the approach to the individual problem. For example, Brown et al.(1991) showed that limits could be set on the form of DEM solutions for OIV and FeXI lines. Methods to deal with and determine the nature of the errors should be investigated. Judge et al. point out that the situation may be improved by using fewer lines in the analysis thus introducing fewer systematic errors or alternatively solving for the systematic errors as well. Alternatively, if the errors appear 'random' Judge et al. point out that the average of multiple inversions helps in the reconstruction. Judge et al.(1995) clearly believe the errors to be systematic as they discount the possibility of variations in abundances as a result. However, it may be more informative to check on this aspect as well, perhaps using the technique of Lang et al.(1990). Finally, the calculations of chap.2 sec.2.1.2 suggest that electron distributions will generally be Maxwellian and that this assumption is valid. However, the reappraisal with respect to the dynamic nature of the atmosphere suggests that departures from Maxwellians should be investigated further.

Of course, despite improving our understanding of the structure of the solar atmosphere over the last twenty years, it is not clear that the DEM method can provide the type of accurate description that will now take us a significant step forward. It tells us something about the volume distribution of emitting plasma globally but the solar atmosphere is continually changing on fine (perhaps unresolved) scales. Observations from SOHO (e.g. Harrison & Insley (1997), Pike & Harrison(1997)) show that bright points, spicules etc. are constantly dynamically evolving and it seems apparent that the idea of a plane-parallel static isothermal atmosphere is outdated. Given the difficulties of advancing the DEM method to two independent variables it may be that predictive modelling is the way forward. Future modelling we wish to attempt is presented as an overview in the next chapter.