THE OPTIMAL EXTRACTION
OF
LOW RESOLUTION IUE SPECTRA:
II. THE NOISE MODEL AND COMPARISONS WITH IUESIPS AND GEX

A. L. Kinney, R.C. Bohlin, and J.D. Neill

Space Telescope Science Institute

June 1988

ABSTRACT

One cornerstone of optimal extraction is rejection of deviant data, such as cosmic ray hits (Horne 1986). In order to distinguish good data from contaminated data, the noise must be accurately modeled. The variance as a function of flux number is derived for the SWP, LWR, and LWP cameras and the difference in this noise for the 55 line files and the 110 line files is discussed for the case of the SWP. Also, a comparison is made between the optimal extraction, IUESIPS, and GEX. Results for a standard star, a bright Seyfert, and a relatively faint AGN are compared.

I. NOISE MODEL

The optimal extraction routine (Kinney and Rivolo 1988, Horne 1986) fits a third order polynomial to the background at each wavelength, excluding pixels that deviate by more than 2.5σ, that have bad data quality flags, or that lie in the 18 (9) lines surrounding the spectrum for the 110 (55) line file. Optimal extraction then normalizes the spectrum to unit integrated flux across the dispersion at each wavelength, averages the normalized flux in each spatial line in bins that can range from a width of 1 to 20 sample points in wavelength, and fits third order polynomials to the bin values along the spectrum for each of the 18 (9) lines, thus, deriving a cross-dispersion profile for each wavelength bin. Finally, the extraction solves for the flux at each wavelength point by fitting the data to the profile at that wavelength. In the final iteration, data is weighted according to its best-fit modeled flux. Data points that differ from the best fit by $n\sigma$ are excluded. (A value of n=5 is used for the test results presented in this paper.) A correct noise model is important to the extraction routine since the σ's are used at every stage to both exclude and to weight the data.

The noise model was determined by measuring the average flux number (and the variance in the flux number (σ_{FN}) within square areas on either side of the spectrum at three different wavelengths, as shown in Figure 1. For the 110 line file, a 26 × 26 pixel region is used; for the 55 line file a 13 × 26 pixel region is used. The images are chosen to cover the full range in FN. Background regions away from the spectrum are used so that a large number of pixels at the same exposure level could be sampled. Twenty-two images are analyzed for the SWP 55 line noise model, sixteen for the SWP 110 line model, and 32 each for the LWP and LWR 55 line model. (The 110 line model has not been constructed yet for the LWP and LWR.) The results are shown in Figures 2 through 5.

For IUE, σ_{FN} is a linear function of FN and can be expressed as

$$\sigma_{FN} = A + B \times FN.$$

(This differs from a CCD, where the noise is dominated by photon statistics, so that σ is proportional to $\sqrt{N}$.) The signal to noise ratio (SNR) is then given by

$$SNR = \frac{FN}{\sigma_{FN}} = \frac{FN}{A + B \times FN},$$

which for large FN is a constant, $\frac{1}{B}$, as might be expected, if the dominant noise is a certain fraction of the signal plus some additional readout noise. The suspicion is that the ITF is not properly removing the fixed pattern that is present at high spatial frequencies in the SEC Vidicon targets of the IUE cameras.

The noise, σ_{FN}, is the greatest at short wavelengths in the SWP and LWP cameras but is unchanged with wavelength for the LWR camera (see also Bohlin *et al.* 1980). Figure 6 shows a comparison between the 55 and 110 line image for the SWP. For a fixed region ,

$$FN_{55} = 2FN_{110},$$

because a pixel from the 55 line file is twice the size of a pixel from the 110 line file. For comparison, the results are plotted with the axis for FN_{110} and σ_{110} expanded by a factor of 2 relative to the parameters for the 55 line file. As expected, $2\sigma_{110}$ is greater than σ_{55} but not as much as $\sqrt{2}$ greater, as should be the case, if all the 110 line pixels are completely independent samples. The resampling of the photometrically corrected image with a spacing of $\frac{\sqrt{2}}{2}$ pixels in both directions in the 110 line file implies some correlation among the adjacent pixels.

II. COMPARISONS BETWEEN OPTIMAL EXTRACTION, IUESIPS, AND GEX

The optimal extraction and the gaussian extraction (GEX) both take into account the shape of the cross-dispersion profile. However, the two algorithms were developed completely independently and use different methods for determining the background, the profile, and the flux at each wavelength. A comparison of these new extractions and the standard IUESIPS for a full range of background and net flux numbers is tested, including a standard star with no radiation background and an optimally exposed spectrum of 10,000 to 40,000 FN, a bright Seyfert galaxy with low background and a good spectrum of 10,000 FN, the same Seyfert in a short exposure with a low background and a poorly exposed spectrum of 3,500 FN, and a faint AGN with high background and a weak spectrum of 3,500 FN. These spectra are HD60753 (SWP28879), NGC5548 (SWP23065), NGC5548 (SWP23066), and 3C390.3 (SWP25523), respectively.

A plot overlaying the three extractions is shown in Figure 7 for the standard star. There are some excursions in features of the order of 5%; but, in general, the spectra track well. Both the optimal extraction and GEX are successful in conserving total flux.

Comparisons are shown in Figures 8 through 11, where the IUESIPS is plotted on the bottom, the optimal extraction in the middle, and the GEX on the top. To quantify the SNR, the point to point deviation has been measured and is expressed as a ratio for optimal extraction or GEX over IUESIPS extraction.

$$R = \frac{\sqrt{\sum (F_i - F_{i+1})^2}_{Opt\ Ext}}{\sqrt{\sum (F_i - F_{i+1})^2}_{IUESIPS}}$$

The point to point deviation is a better measure of the noise properties than the standard deviation because continuum features like a moderate slope do not dominate.

However the point to point deviation will increase with the presence of emission lines, and so should be used in regions absent of features. Table 1 summarizes R for S28879, S23065, S23066, and S25523 in various 100Å bins devoid of emission lines.

As expected, there is little or no improvement when well exposed spectra, such as the standard star (S28879) and the bright Seyfert (S23066), are optimally extracted. However, the extraction becomes important for exposures of lower FN. This improvement is clear in Figure 10 for SWP23066, *where the ratio of point-to-point deviation is 0.65 for both the optimal and GEX extractions.* The result is less encouraging for S25523, where GEX shows a lower point-to-point deviation than optimal extraction but does not succeed in removing the cosmic ray at 1250Å. Since the two algorithms are so dissimilar, the source of the lower point-to-point deviation of GEX is difficult to trace. One possibility is the background subtraction, since S25523 is the image with the highest background. This image requires more analysis.

III. CONCLUSIONS

The noise (σ_{FN}) varies linearly with flux number (FN). Sigma decreases with increasing wavelength in the SWP and LWP cameras but is unchanged with wavelength in the LWR camera.

Both the optimal extraction and GEX dramatically improve the point-to-point deviation in spectra that are not optimally exposed, while conserving total flux. Currently, optimal extraction is better at dealing with cosmic ray rejection; but there is at least one case where GEX is more successful at reducing point-to-point variation. Further analysis and comparisons must be done to learn whether either of these algorithms could be used in a production reprocessing of all IUE low dispersion spectra.

TABLE 1

Ratio R of Deviation of Optimal Extraction and GEX to IUESIPS

IMAGE	WAVELENGTH	OPTIMAL	GEX
SWP28879	1700-1800Å	1.06	1.10
	1800-1900Å	1.04	1.00
SWP23065	1700-1800Å	0.80	0.98[a]
	1800-1900Å	0.83	0.86
SWP23066	1700-1800Å	0.67	0.65
	1800-1900Å	0.65	0.67
SWP25523	1400-1500Å	0.94	0.79
	1700-1800Å	0.99	0.73
	1800-1900Å	0.99	0.74

(a) GEX does not currently remove the reseau that lies in the spectral region at about 1790Å (cf. Figure 7).

References

Bohlin, R.C., Holm, A.V., Savage, B.D., Snijders, M.A.J., and Sparks, W.M. 1980, *A.A.*, **85**, 1.

Horne, K. 1986, *P.A.S.P.*, **98**, 609.

Kinney, A.L., Riviolo, A.R. 1988, in *IUE Newsletter*, No. 34, p. ...

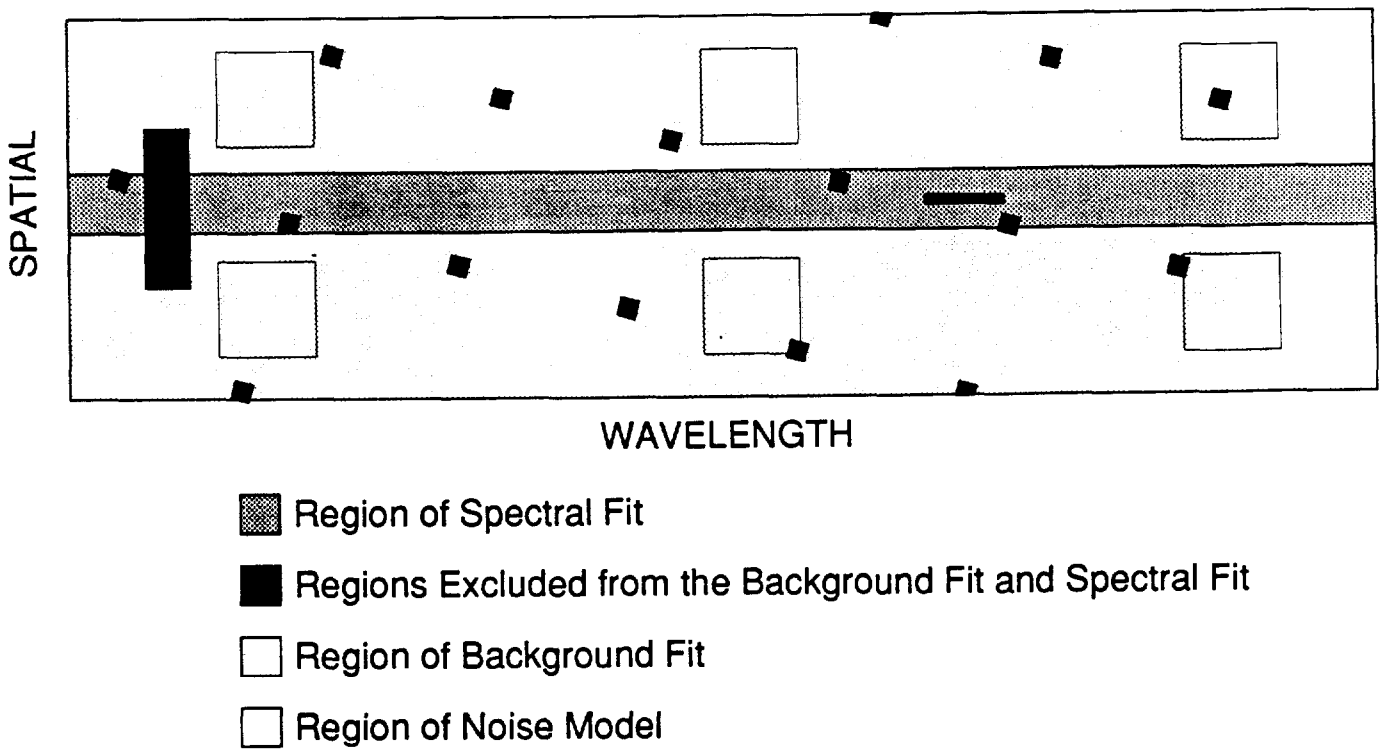
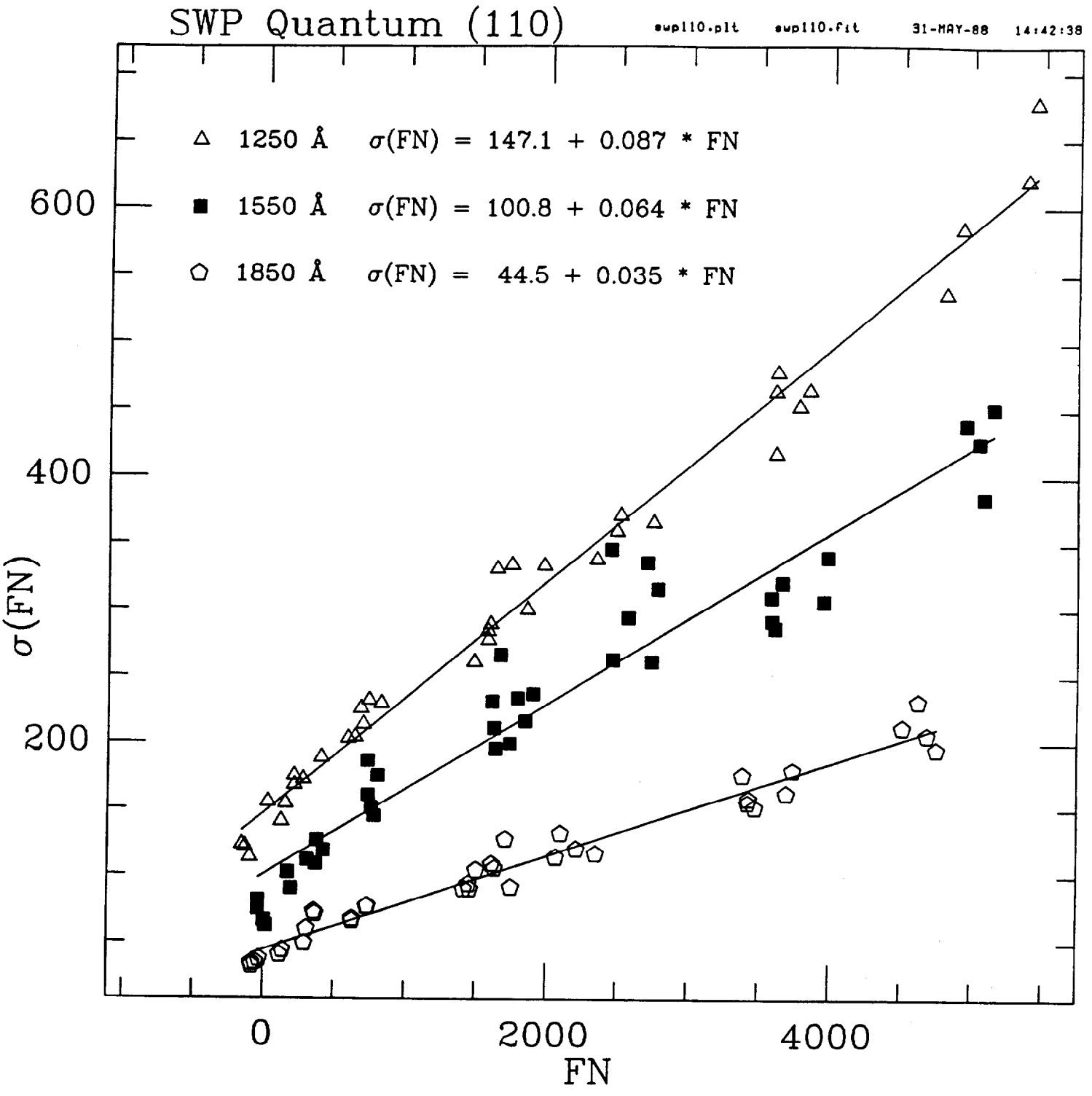


Figure 1. A schematic of the line-by-line file showing the regions used for measuring the noise model. The regions are 26 × 26 (13 × 26) pixels and are placed with the edge nearest to the spectrum being 16 (8) pixels from the spectrum center in the 110 (55) line files, respectively.


SWP Quantum (110)
swp110.plt
swp110.fit
31-MAY-88
14:42:38
1250 Å σ(FN) = 147.1 + 0.087 * FN
1550 Å σ(FN) = 100.8 + 0.064 * FN
1850 Å σ(FN) = 44.5 + 0.035 * FN
600
400
200
σ(FN)
0
2000
4000
FN

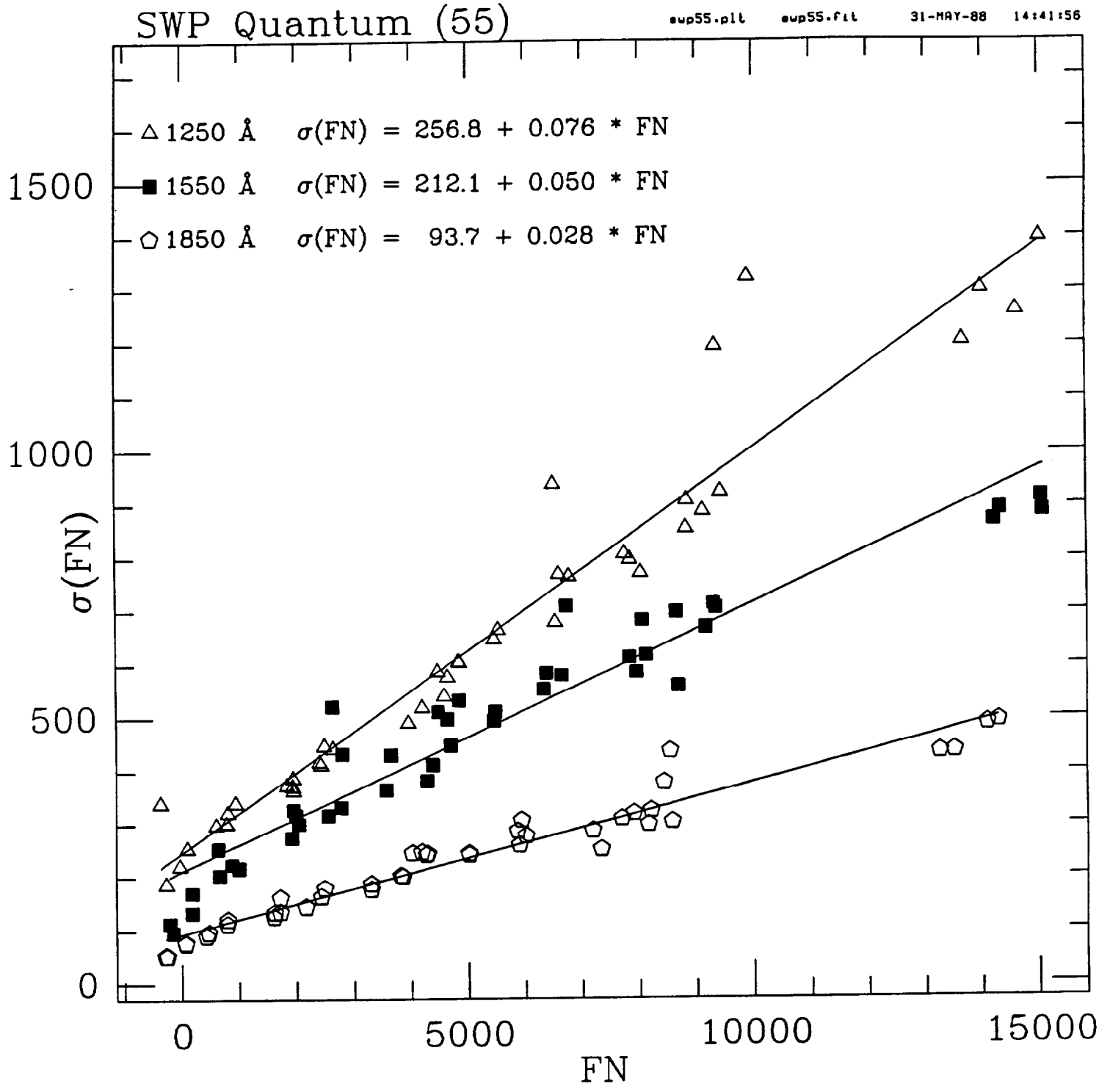


Figure 3. Noise model for the SWP 55 line file.

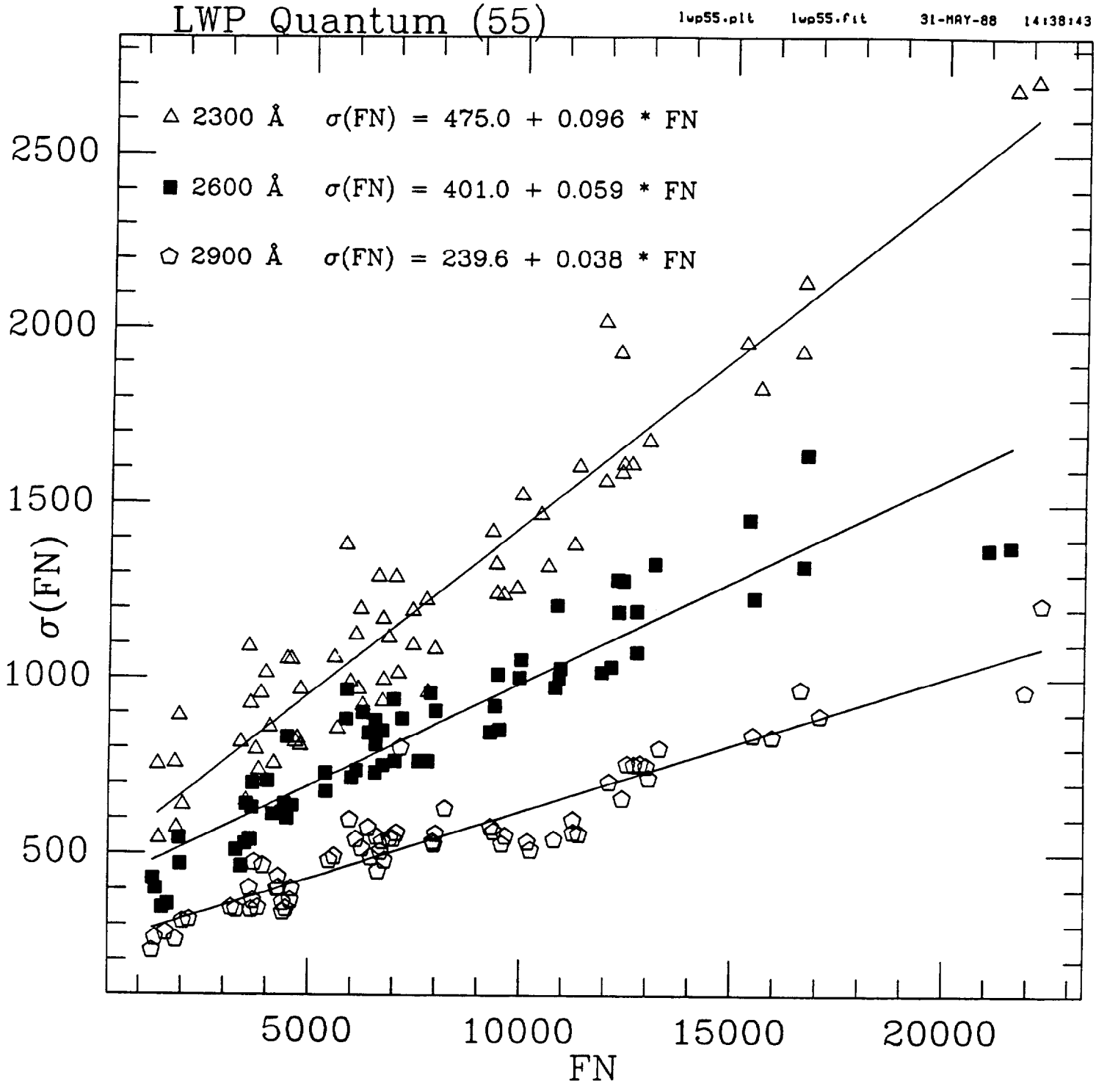


Figure 4. Noise model for the LWP 55 line file.

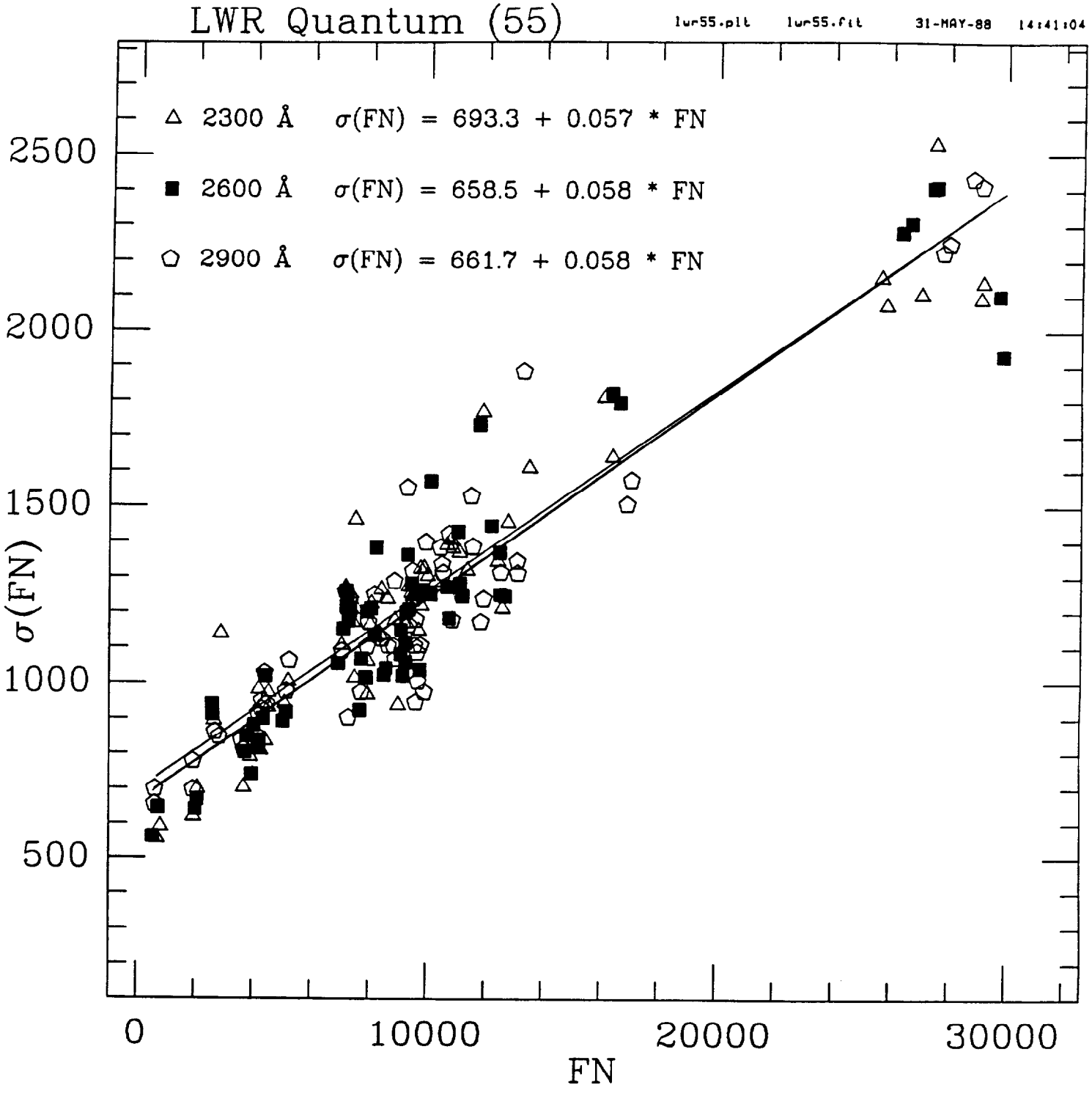


Figure 5. Noise model for the LWR 55 line file.

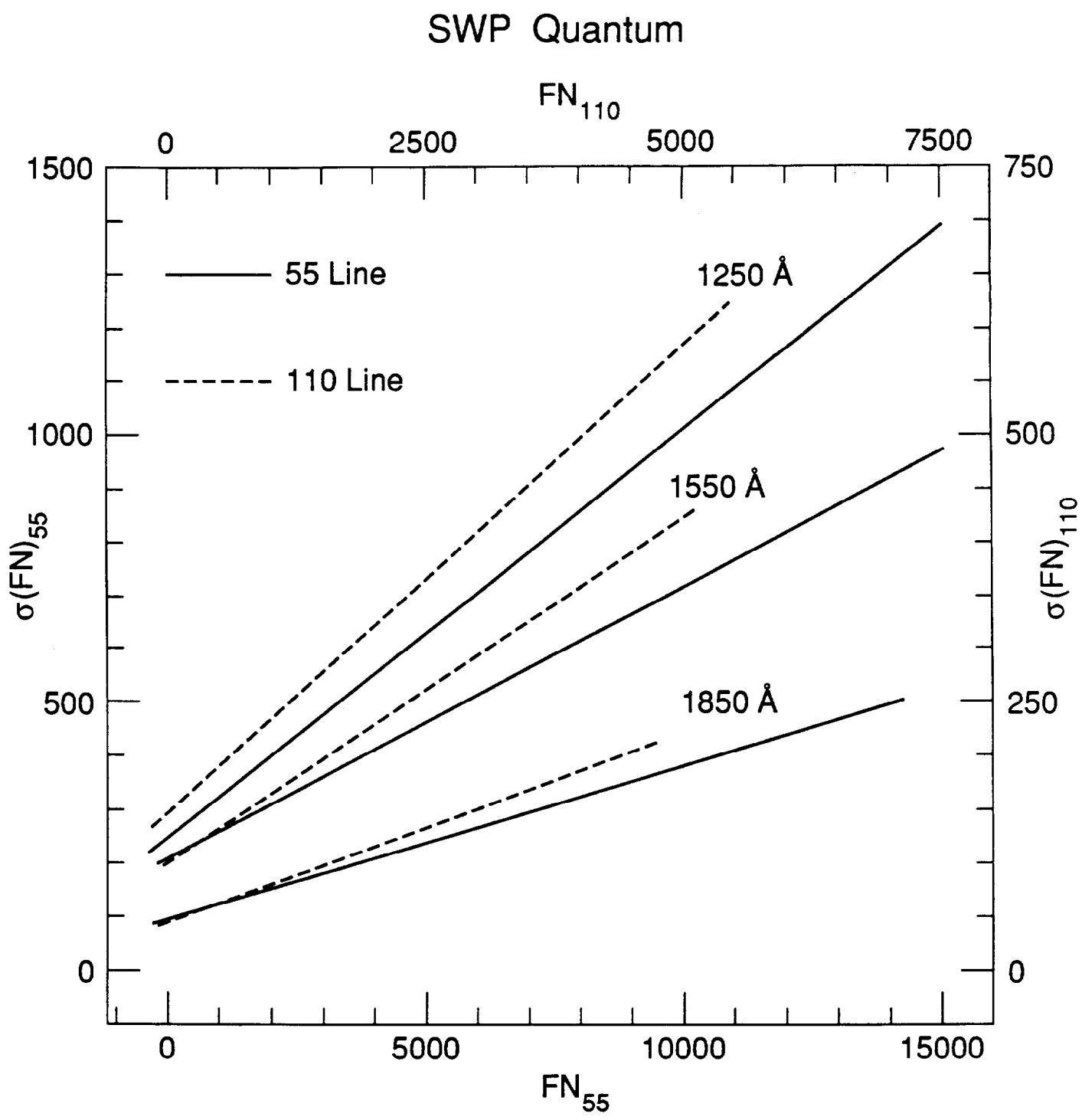


Figure 6. Comparison between the 55 and 110 line file for the SWP.

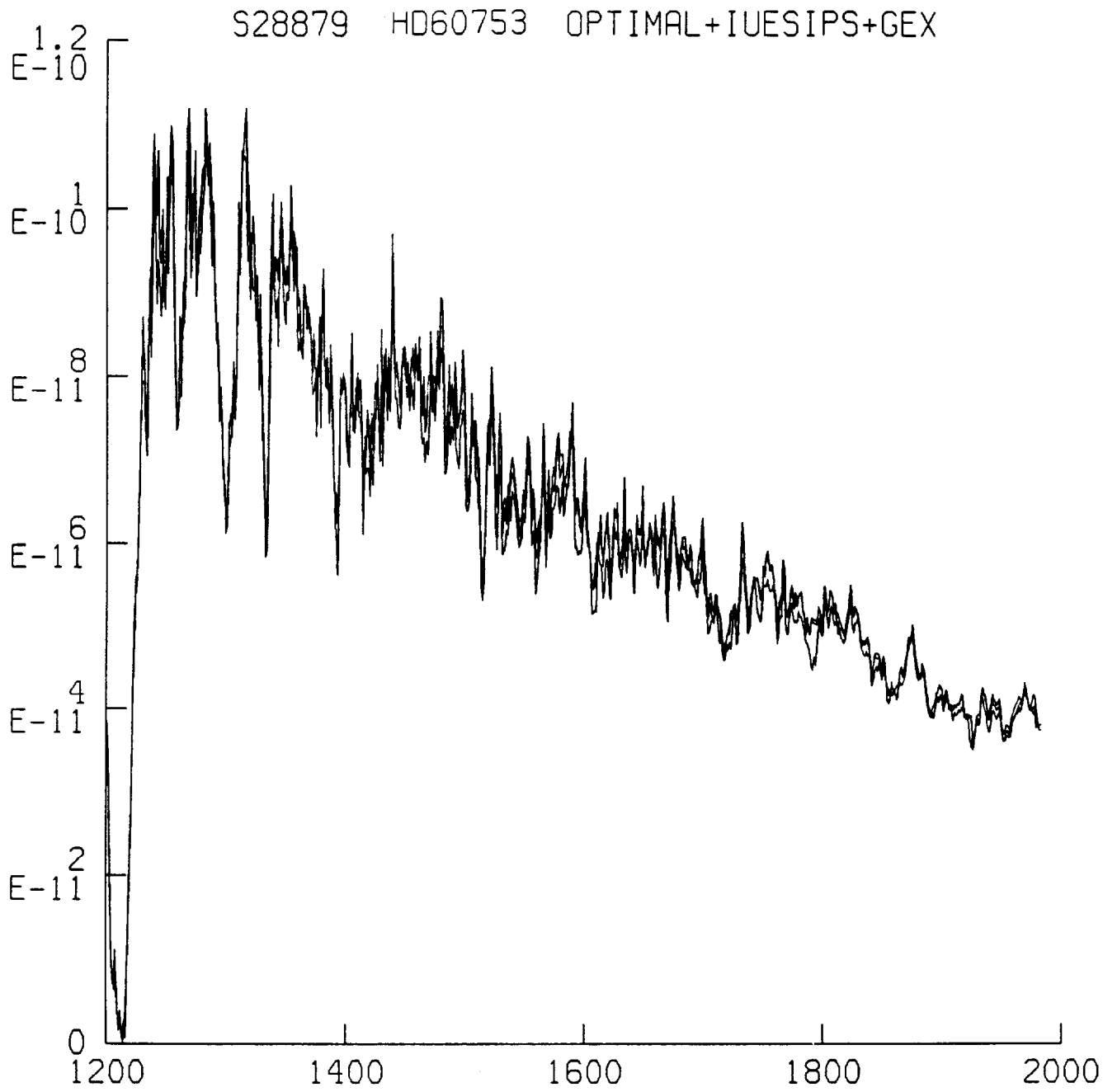


Figure 7. The three extractions are overlaid for the standard star HD60753. Although there are some excursions in features of the order of 5%, in general the spectra track well.

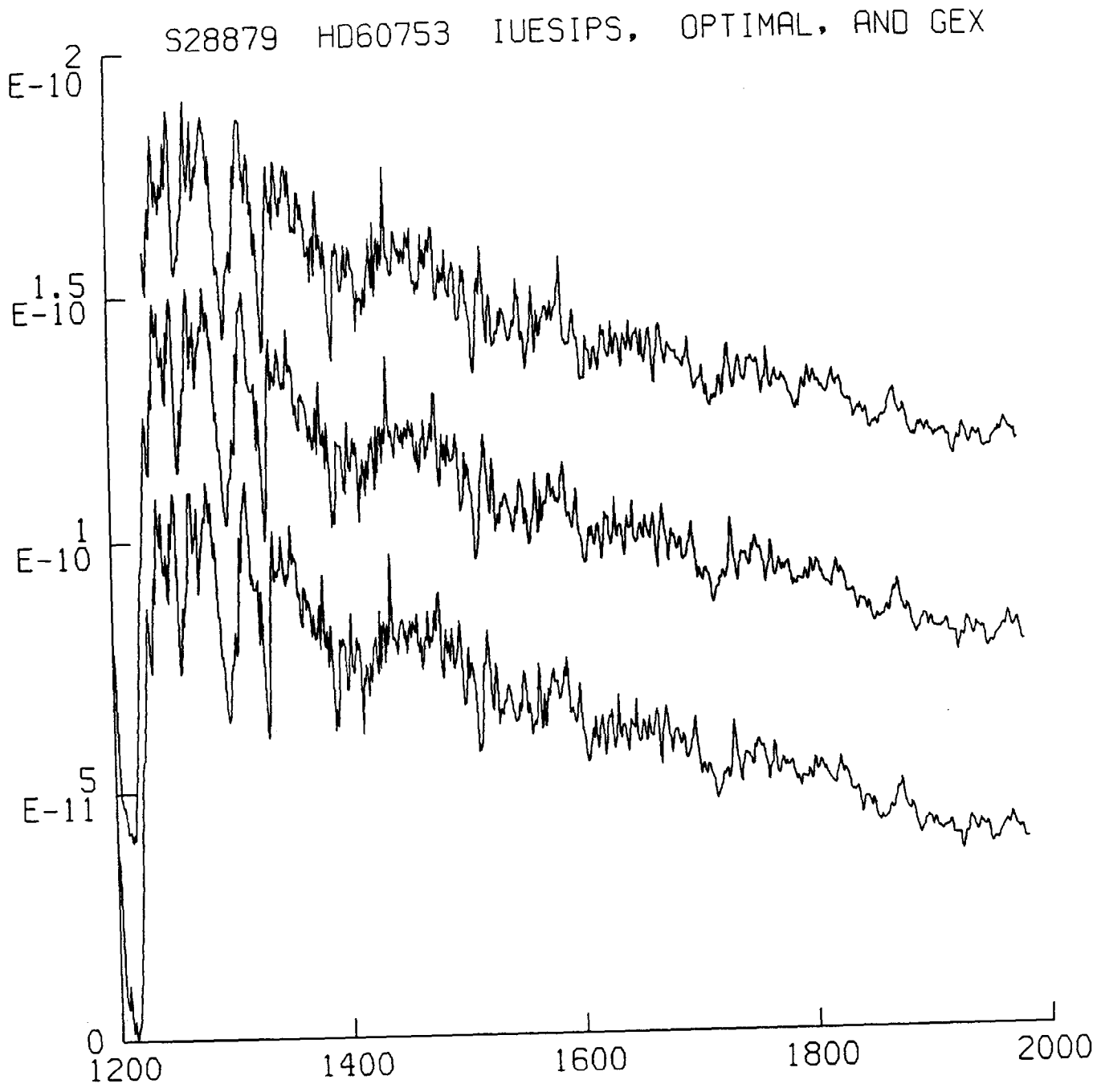


Figure 8. The three extractions of SWP28879 are plotted with IUESIPS on the bottom, with optimal extraction offset by $4. \times 10^{-11}$, and with GEX extraction offset by $8. \times 10^{-11}$.

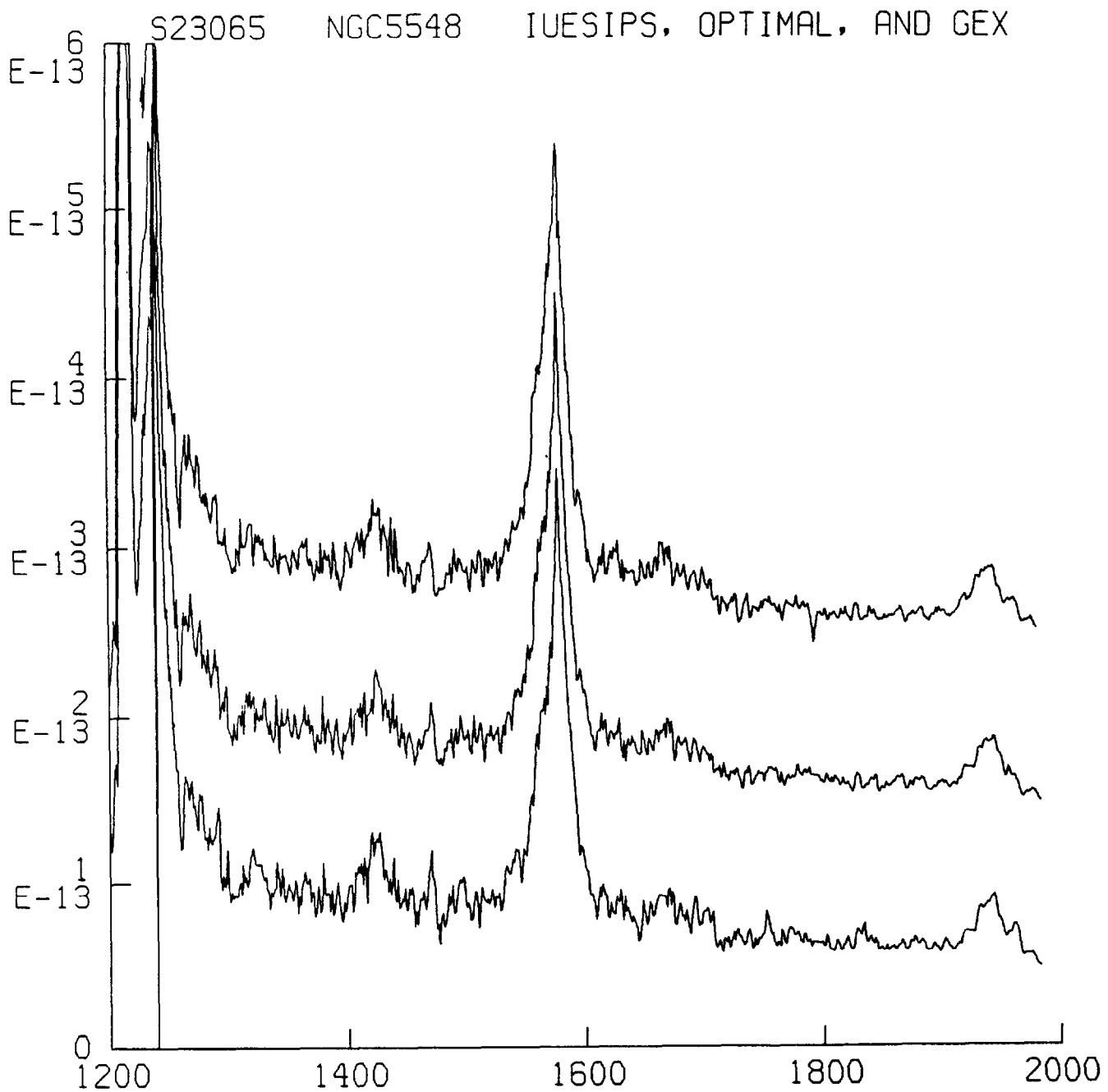


Figure 9. SWP23065 is plotted with IUESIPS on the bottom, with optimal extraction offset by $1. \times 10^{-13}$, and with GEX extraction offset by $2. \times 10^{-13}$.

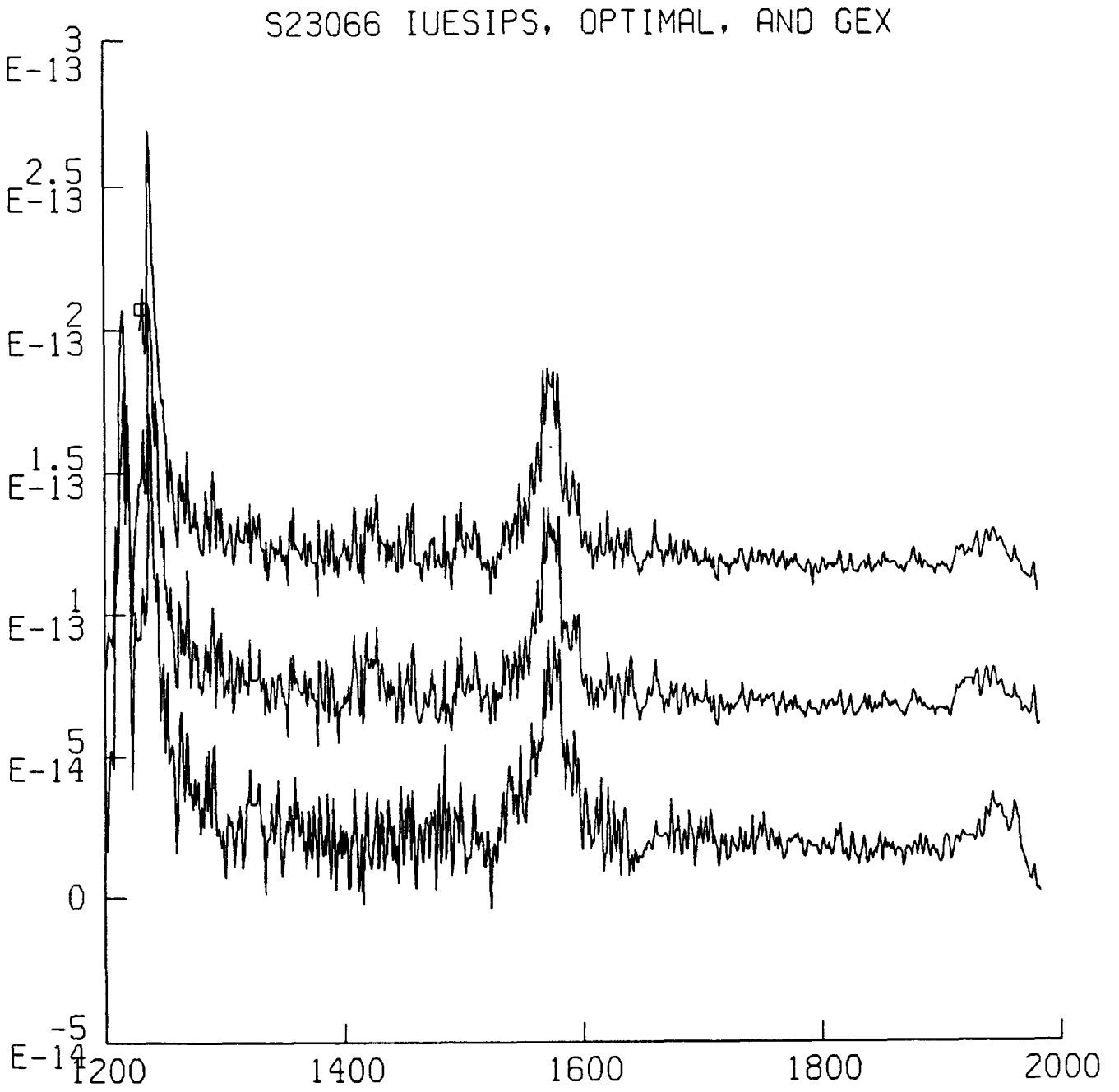


Figure 10. SWP23066 is plotted with IUESIPS on the bottom, with optimal extraction offset by $5. \times 10^{-14}$, and with GEX extraction offset by $1. \times 10^{-13}$.

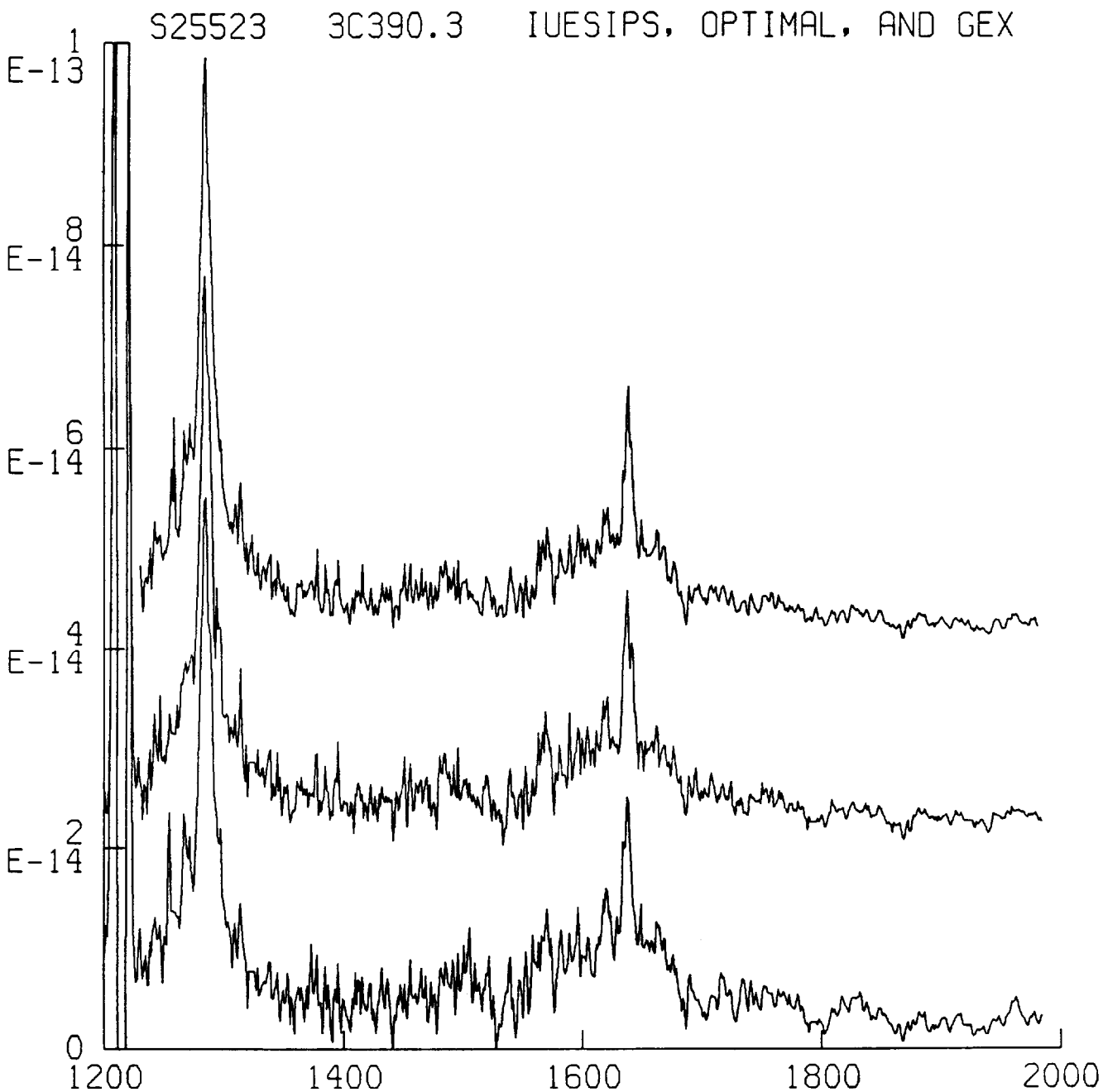


Figure 11. SWP25523 is plotted with IUESIPS on the bottom, with optimal extraction offset by 2. $\times 10^{-14}$, and with GEX extraction offset by 4. $\times 10^{-14}$.