

## Some Neglected Pixel Problems

Kris Davidson

*School of Physics and Astronomy, University of Minnesota, 116 Church St. SE,  
Minneapolis MN 55455*

**Abstract.** We have developed several improved techniques because standard reduction software was inadequate for the Treasury Project on  $\eta$  Carinae. Though adapted specifically to HST/STIS/CCD data, some of these methods can be useful for images obtained with other instruments. For instance, our subpixel modeling technique reduces some of the evil effects of marginal sampling, i.e., in images where the pixels are moderately too large to take full advantage of the instrument's basic p.s.f. Marginal sampling causes the apparent p.s.f. to depend on precise location in the pixel grid, a subtle effect that perturbs measurements. Our method considerably reduces the amplitude of this phenomenon.

### 1. Preliminary Comment: Not Just Spectroscopy

The techniques sketched here were motivated by intensive STIS spectroscopy of one object. But their usefulness is far more general; they can be applied to *a broad class of images obtained with various instruments*. Whenever the p.s.f.'s FWHM is between about 1 and 3 pixel widths, one should consider "subpixel modeling" as outlined in Section 4.

### 2. The Role of $\eta$ Carinae Concerning HST Data in General

Before getting into data processing techniques, let me frankly indulge in some advertising which serves a valid purpose here. Most astronomers have a nodding acquaintance with  $\eta$  Car: a very massive star with intense, ostentatiously bipolar ejecta. The degree of structural complexity, however, and the surprisingly basic unsolved problems lurking there, are less well known.<sup>1</sup> Together these characteristics have made Eta unique in HST's record:

- The ejecta are *spatially complex* on size scales of 0.01 to 10 arcsec;
- They and the star are *spectrally complex* too, with thousands of identified emission lines;
- At least six distinct types of spectra are found there, relevant to diverse problems;
- And some of those spectra vary systematically in time.
- Thus we have needed excellent spatial and spectral resolution, appreciable spatial coverage, broad wavelength coverage, and repeated observations, all together.
- Both the star and its ejecta are *bright enough to make such observations feasible* without requiring huge amounts of telescope time.

Scarcely any other major HST target shares this combination of attributes. As a result,  $\eta$  Car has consistently provided unusual illustrations of HST's capabilities. (Why this matters to us here, will become evident in the next paragraph.) Since 1991 this object has held the record for angular resolution in spectroscopy of a complex target, first with

---

<sup>1</sup>For general information about this object, see Davidson & Humphreys (1997) and three proceedings volumes: Humphreys & Stanek (2005), Gull et al. (2001), and Morse et al. (1999). The HST Treasury Program for Eta has produced numerous scientific results, but here we're concerned with data techniques.

the FOS, then GHRS, and, since 1998, with the STIS. I honestly can't think of another object that demonstrated the STIS as well – for instance, almost nothing else was observed across the entire wavelength range, as we did repeatedly for Eta. Several HST instruments have been used *intensively* on this object, close to their resolution limits, sometimes with unconventional techniques. The 72-orbit Treasury Project obtained a vast number of meaningful data pixels (more than a typical multi-hundred-orbit extragalactic program, for instance), because many integrations could be obtained in each orbit and because of the extended structure.<sup>2</sup> In some respects this was the most ambitious spectroscopic program ever attempted with the HST.

Standard reduction software and standard information about the instrument proved to be inadequate for such intensive observations. Therefore, entirely apart from scientific goals, *these data were and are valuable in three practical ways*: (1) They clearly and quantitatively revealed various glitches in the STIS/CCD and in the data reduction procedures; (2) they forced us to develop techniques to reduce those problems; and (3) they provide a unique, public archival record of instrument characteristics as functions of time. The entire set of STIS data on  $\eta$  Car, including pre-project observations incorporated into the overall data set, extend from 1998.0 to 2004.2, most of the instrument's working life.<sup>3</sup>

We hope these facts may be useful for improved analyses of existing STIS data on other objects – particularly those where the spatial resolution proved to be unsatisfying because of shortcomings in the reduction software, not the instrument itself (see below). Some of the techniques can be applied to other instruments as mentioned earlier.

### 3. A Brief List of Glitches

Thus, as foreseen in the original proposal, a substantial fraction of our effort in the  $\eta$  Car Treasury Program has gone into improving the STIS/CCD data processing methods and assembling a convenient, somewhat unusual data archive. In this section I'll just list the main difficulties we've dealt with; here "we" means mainly John Martin, Kazunori Ishibashi, and myself, the persons working on these development tasks.<sup>4</sup> Additional information on the following effects and techniques can be found at the project website, <http://etacar.umn.edu>.

(1) At UV to visual wavelengths the STIS CCD provided only marginal spatial sampling – i.e., the 0.05 arcsec pixel size was undesirably large relative to the p.s.f. width. With standard software, *this fact severely degraded the spatial resolution of spectroscopy*. The reason, which concerns variations in local sampling, is more complicated than one might assume at first glance. In order to get resolution better than 0.2 arcsec with this instrument, one must extract a spectrum less than 4 CCD rows wide. Since a spectral "trace" is slightly tilted relative to a CCD row, the spatial sampling varies as a function of wavelength; at some detector columns the spectrum of a point source coincides with a row center, while at others it falls midway between two rows. With normal interpolation methods, this systematic variation in sampling causes any narrow continuum extraction to appear "scalloped" or wavy. If one attempts to observe a target near another object, the contamination cannot accurately be subtracted out – at least, not with standard software – because its undulations

---

<sup>2</sup>Of course 72 orbits is a big number by most users' standards. But those orbits were required to study intricate variations applicable to several branches of astrophysics, not just a basic set of spectra.

<sup>3</sup>One potentially useful detail: Certain ejecta blobs have rich spectra of narrow, well-classified emission lines, which provide supplementary wavelength calibrations.

<sup>4</sup>For some purposes Matt Gray, Michael Koppelman, J. T. Olds, and R. M. Humphreys should also be included in the list. Other project CoI's have been concerned mainly with various aspects of the scientific results, as planned.

generally differ in both phase and amplitude from those of the target object. We suspect that this phenomenon discouraged some early STIS users from making further attempts to do high-spatial-resolution spectroscopy. Dithering could have improved the sampling but was not always feasible – for instance, we didn’t have enough telescope time to dither the Treasury Project spectroscopy. Fortunately, a technique sketched in Section 4 below – essentially a novel form of pixel interpolation – provides a considerable improvement. This method can be applied to existing STIS data and to other image data whose pixels are somewhat larger than one would have preferred.

(2) Because our subpixel modeling technique (effect 1 above and Section 4 below) depends on local arrays of about 20 pixels, it was necessary to eliminate practically all bad pixels – hot pixels, cosmic ray hits, etc. Therefore we employed a bad-pixel identification method in addition to the normal CR-SPLIT procedure. To test each pixel, our software analyzed the dispersion of pixel values in a  $5 \times 5$  sample array centered at that location. This technique works rather well even if only one image is available, i.e., if no CR-SPLIT companion image exists.

(3) The complete STIS/CCD wavelength range, UV to far red, required about 30 grating tilts. In order to produce a unified spectral image covering the entire range, we need to splice each adjoining pair of images. However, there is an unexpected complication: the detector was not uniformly focussed. In fact the p.s.f. was generally sharper at the longer-wavelength side of each image. Therefore we have had to develop a convolution procedure to make adjacent wavelength samples match in their overlap region. This is further complicated because tiny variations in pointing caused the apparent star/ejecta brightness ratio to vary somewhat.

(4) Each STIS/CCD spectral image contains several types of “ghosts” with various intensities and size scales. We have parametrized the worst of these, especially those with small size scales. With this information they can be reduced by a simple procedure that basically amounts to Jansson – Van Cittert deconvolution.

(5) The STIS/CCD spatial p.s.f. had asymmetric wings. We have measured the parameters necessary to reduce this effect, in much the same manner as effect 4 above.

(6) Suitable ERR arrays – local noise levels stored in each FITS file – are not entirely obvious. In particular, when the stored pixel values have been produced by interpolation between original data pixels (see Section 4 below), then one should *not* store the mathematically correct r.m.s. error for each interpolated pixel. Reason: The interpolation process causes adjacent pixel values to be correlated with each other. In that case one cannot estimate the true error of a sum of pixels merely by taking a quadratic sum of the individual formal errors. – I recognize that these comments may seem a little cryptic, but a full explanation would take too much space here so the reader is invited to examine the problem himself/herself. The point is that the ERR arrays in some pipeline-processed STIS images are misleading, and don’t contain enough information to reconstruct the true error in a sum or average of many pixels. In fact an uncritical application of them will usually lead to an underestimate of the noise level, sometimes a serious underestimate.

There are other difficulties but these are probably the most troublesome. Items 1, 2, and 6 apply to image data in general, and effect 3 might also have applications beyond STIS. The remainder of this paper concerns item 1.

#### 4. Optimum Subpixel Modeling

For quantitative work we usually need to correct optical distortions, to rotate the image, to make position measurements, etc. These tasks require some form of interpolation between the original pixel locations. Often this “subpixel modeling” is implicit or hidden in the software, but it is essential; and marginal sampling makes it trickier.

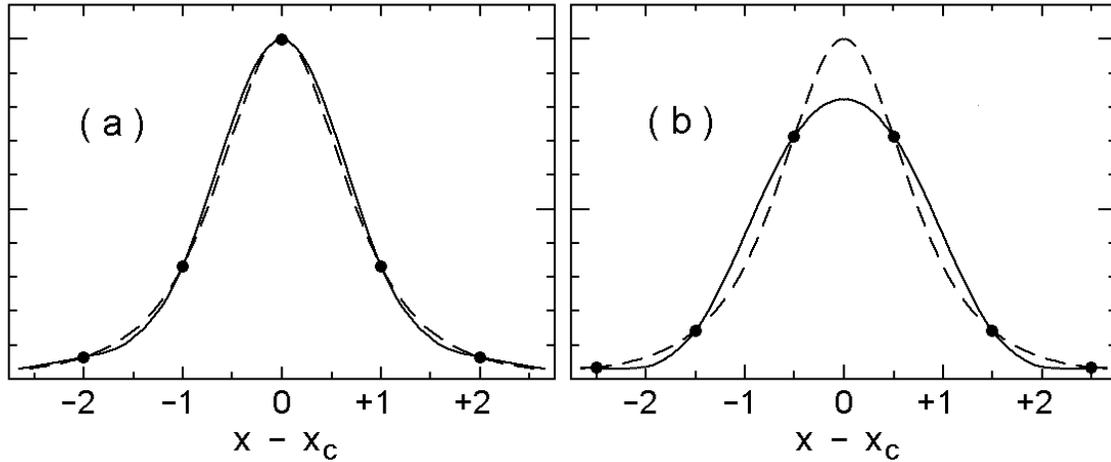


Figure 1.: The dashed curves represent a function with  $\text{FWHM} = 1.5$  pixels. The continuous curves are spline fits to the sample points shown as dots. Here  $f(x)$  is the square of a Lorentzian,  $[1 + \alpha(x - x_c)^2]^{-2}$ , which closely resembles the p.s.f.'s found in STIS/CCD data.

By “marginal sampling” I mean any case where the p.s.f. width (FWHM) is between, roughly, 1 and 3 pixel widths. Sampling naturally degrades the average resolution, but that isn't the main problem. Explanations are easier if we idealize the situation in two ways: (1) Rather than a two-dimensional image  $f(x, y)$ , let's consider the one-dimensional case  $f(x)$ . Our unit of  $x$  will be the original pixel width. (2) In a real image each pixel value is some sort of average across the pixel, but it's easier to discuss uniformly spaced samples of a continuous function,  $f_i = f(x_i)$ . Neither of these simplifications affects the basic problem; the approach described below can be generalized to real data.

Figure 1 shows a function – maybe a p.s.f. – sampled with two different pixel offsets. In Fig. 1a the peak coincides with a data point (a pixel center), and a cubic spline through the data points matches the function pretty well. In Fig. 1b, however, the peak falls midway between two sample points; then the spline fit is much worse, with a severely depressed peak and a broadened profile. Evidently *the effective p.s.f. depends on a point source's location on the detector*. This fact is fundamental, not merely a consequence of the interpolation scheme. The sampling is blind to some information near and beyond the critical Fourier frequency  $\omega_c = \pi \text{ pixel}^{-1}$ , and *the precise location of the pixel grid determines which Fourier components are lost*. One doesn't notice the resulting spatial fluctuation of the p.s.f. when viewing an image, but it affects measurements. This phenomenon is responsible for the undulations in narrow spectral extractions, effect 1 in Section 3 above. Incidentally, one can show that deconvolution methods don't help this problem much.

In principle the effective sampling frequency can be improved by “dithering,” taking multiple exposures with fractional-pixel relative offsets. But dithering requires extra effort and can be inefficient in telescope usage; it depends on reliable, extremely precise offset pointing; it can fail for various reasons without the failure being obvious; and sometimes one simply doesn't have enough telescope time. A *consistent* resolution of about 2 pixels (e.g., 0.1 arcsec with the STIS/CCD) can be attained without dithering, if we use a special technique outlined below. This subpixel modeling technique greatly reduces the spatial variation of the p.s.f.

Here's our plan. We use the original data points to generate a related function  $g(x)$  with the following goals:

1.  $g(x)$  approximates  $f(x)$  as closely as possible without excessively violating requirement 2 below. Function  $g(x)$  will turn out to be a blurred version of  $f(x)$  and of course we hope to minimize the blurring.
2. A second goal, motivated by Fig. 1, distinguishes our effort from ordinary interpolation: *The calculated shape of  $g(x)$  must depend as little as possible on the location of  $f(x)$  relative to the pixel grid.* For instance, imagine two images  $f_1(x)$  and  $f_2(x)$  with the same shape but relatively offset by half a pixel:  $f_2(x) = f_1(x - 0.5)$ . We design the corresponding  $g$ -functions so that  $g_2(x) \approx g_1(x - 0.5)$  as accurately as feasible. As Fig. 1 implies, this requirement is subtle. It isn't easy to achieve without heavy blurring.
3. Based on  $f(x_i)$  at integer points  $x_i = i$ , we calculate  $g(x_k)$  with half-integer spacing,  $x_k = k/2$ . Since this reduced interval provides decent sampling, we can later estimate  $g(x)$  at any other point by conventional interpolation.

Goals 1 and 2 tend to conflict with each other, so the “best” solution for any given data set is a compromise. Moreover, since quantitative criteria depend on the specific application, there can be no universal solution. Our reasoning will be fairly general.

Since the procedure for generating  $g(x)$  must be linear in  $f(x)$ , each output sample point  $g(x_k)$  must be a linear combination of nearby input data  $f(x_i)$ . We further assume symmetry between the positive and negative  $x$ -directions; this is not strictly true for some instruments (including STIS) but it's a good practical approximation that greatly simplifies the effort. These two requirements determine the formula for locations that coincide with the original data points  $x_i$ . If  $x_k$  is an integer, then

$$g(x_k) = A_0 f(x_k) + \sum_{n=1}^{M-1} A_n \cdot \{ f(x_k - n) + f(x_k + n) \} , \quad (1)$$

where the  $A_n$  are constant coefficients. Similarly, for the intermediate output points where  $x_k = \text{integer} + 1/2$ ,

$$g(x_k) = \sum_{n=0}^{N-1} B_n \cdot \{ f(x_k - 0.5 - n) + f(x_k + 0.5 + n) \} . \quad (2)$$

In a practical scheme  $M$  and  $N$  will usually be 2, 3, or 4. “Conservation of counts” requires

$$A_0 + 2 \sum_{n=1}^{M-1} A_n = 2 \sum_{n=0}^{N-1} B_n = 1 . \quad (3)$$

Moreover, if

$$\sum_{n=1}^{M-1} n^2 A_n = \sum_{n=0}^{N-1} (2n+1)^2 B_n = 0 , \quad (4)$$

then  $g(x)$  exactly matches  $f(x)$  if the latter is any cubic polynomial. Constraint 3 is essential but 4 is not.

Let me emphasize that *any* linear symmetric scheme will be equivalent to eqns. 1 and 2, even if it's presented in terms of Fourier analysis or other math language.<sup>5</sup>  $|A_{n+1}|$  tends to be much smaller than  $|A_n|$  and formula 1 either blurs or sharpens  $f(x)$ , depending on

---

<sup>5</sup>If we drop the assumption of symmetry in the  $x$ -direction, then we still have linear combinations of  $f(x_i)$ , but with almost twice as many coefficients.

whether  $A_1$  is positive or negative. For reasons that I don't have space to prove here, we must accept the former, the blurred case. Formula 2 is like standard interpolation but the coefficients are based on non-standard criteria. If we later generalize to realistic pixels, i.e., local averages of  $f(x)$ , only the coefficient values will be altered.

The coefficients  $A_n$  and  $B_n$  will be based on goals 1 and 2 stated earlier. A criterion for goal 2 can be defined as follows. First, choose a test function  $f(z)$  like that shown in Fig. 1; here we'll use coordinate  $z = x - x_c$  instead of  $x$ , where  $x_c$  is some arbitrary offset. The test function peaks at  $z = 0$ , it's normalized so that  $f(0) = 1$ , and it resembles the instrumental p.s.f. Consider any particular value of  $z$ . If we sample the function at  $z, z \pm 1, z \pm 2$ , etc., then a  $g$ -value can be calculated from eqn. 1:

$$g_A(z) = A_0 f(z) + A_1 \cdot \{f(z-1) + f(z+1)\} + \dots$$

If, instead, we sample it at  $z \pm 0.5, z \pm 1.5$ , etc., then eqn. 2 becomes applicable:

$$g_B(z) = B_0 \cdot \{f(z-0.5) + f(z+0.5)\} + B_1 \cdot \{f(z-1.5) + f(z+1.5)\} + \dots$$

The difference  $g_B(z) - g_A(z)$  indicates the local departure from requirement 2. A reasonable measure of the overall discrepancy is

$$E_{AB} = \left[ \int_{-\infty}^{+\infty} \{g_B(z) - g_A(z)\}^2 dz \right]^{1/2}. \quad (5)$$

We adjust the coefficient sets  $A_n$  and  $B_n$  to minimize  $E_{AB}$ , subject to other requirements. Parameter space must be explored with many calculations – at least I haven't thought of any other way to do it. For some applications, it may be useful to emphasize selected parts of the p.s.f. by means of a weighting function  $w(z)$  in the above integral.

Fourier considerations show why there is no really satisfying set of coefficients. If the test function  $f(z)$  is symmetric, then we need only the cosine transforms  $F(\omega)$ ,  $G_A(\omega)$ , and  $G_B(\omega)$ , referring to  $f(z)$ ,  $g_A(z)$ , and  $g_B(z)$ . Then formulae 1 and 2 have the following effects:

$$G_A(\omega) = H_A(\omega) F(\omega) \quad \text{and} \quad G_B(\omega) = H_B(\omega) F(\omega),$$

where

$$H_A(\omega) = A_0 + 2A_1 \cos \omega + 2A_2 \cos 2\omega + \dots,$$

$$H_B(\omega) = 2B_0 \cos 0.5\omega + 2B_1 \cos 1.5\omega + \dots$$

$H_A(\omega)$  and  $H_B(\omega)$  are filtering functions that dampen Fourier components near the critical frequency  $\omega_c = \pi$ , and the parameter  $E_{AB}$  (eqn. 5) can be expressed in terms of a frequency integral of  $\{(H_B - H_A) F\}^2$ . Our "goal 1" stated earlier requires a rapid cutoff as close as possible to  $\omega_c$ , while goal 2 requires  $H_A(\omega) \approx H_B(\omega)$ . Unfortunately, as Fig. 2 (next page) shows, *the two filtering functions are largely incompatible at high frequencies* – because  $dH_A/d\omega = 0$  and  $H_B = 0$  at  $\omega = \omega_c$ , two very different conditions. Moreover,  $H_A$  and  $H_B$  have opposite slopes above the critical frequency. Since these are fundamental consequences of the sampling, goals 1 and 2 cannot both be satisfied to high accuracy.

Finding a suitable compromise is difficult. We can force  $H_A$  to  $H_B$  resemble each other for  $\omega < \omega_c$  by imposing the additional constraints  $H_A = 0$  and  $dH_B/d\omega = 0$  at  $\omega_c$ . Then, however, we find that the cutoff occurs at lower-than-desired frequencies, which means that  $g(x)$  is severely blurred compared to the test function  $f(z)$ . In order to lessen the blurring, we must accept a set of coefficients that causes  $H_A$  to dip slightly below zero at  $\omega_c$  where  $H_B$  automatically passes through zero. This gives acceptable results provided that the cosine transform  $F(\omega)$  is small at the critical frequency – which is a fancy way of saying that  $f(z)$  must not be too narrow.

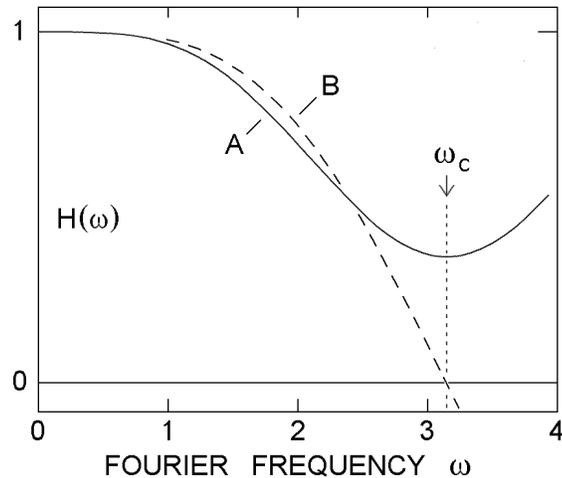


Figure 2.: General shape of the filtering functions  $H_A(\omega)$  and  $H_B(\omega)$  for typical sets of coefficients. At the critical frequency,  $H_A$  is level while  $H_B$  passes through zero. (The particular coefficients used here do not give a useful solution to our problem, but were chosen merely to illustrate behavior near  $\omega_c$ .)

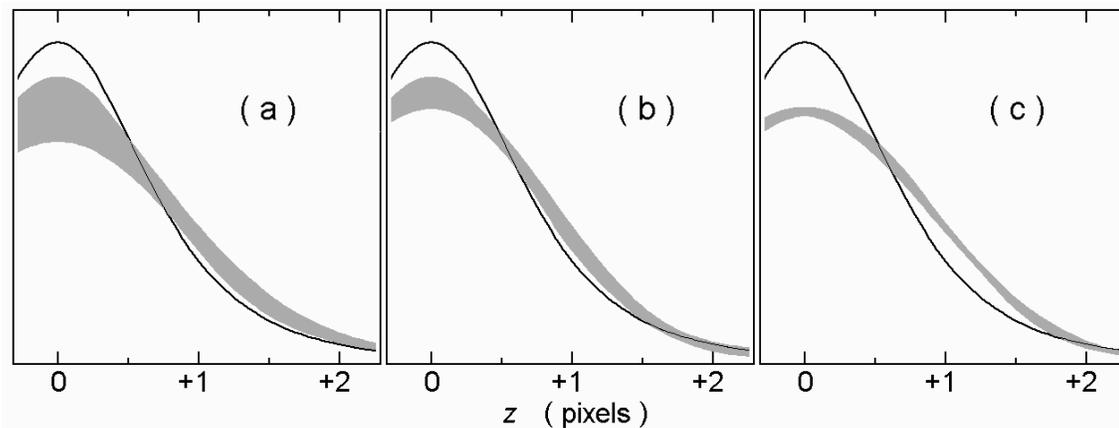


Figure 3.: Interpolation envelopes for (a) linear interpolation, (b) spline interpolation, and (c) eqns. 1 and 2 with coefficients mentioned in the text. In each case the solid curve represents the instrumental p.s.f. before it is sampled by the detector. After sampling and interpolation, the effective p.s.f. depends on the precise subpixel location of the point source (compare Fig. 1); gray areas here represent the envelopes of interpolated values for all possible pixel-grid offsets.

In fact the necessary blurring doesn't hurt as much as one might imagine. Referring back to Fig. 1, the effective output p.s.f. will be a little broader than the solid curve of case 1b. This *appears* much worse than 1a. But the resolution shown in Fig. 1a is largely illusory; it depends on placing the pixel grid in an advantageous position and then knowing that this has been done. Fig. 1b gives a more realistic impression of the true resolution for most applications. I emphasize again that any linear symmetric procedure will be basically equivalent to this one, even if it's described in different words and equations.

Here I won't recommend specific  $A_n$ ,  $B_n$  coefficient values, for three reasons: Several alternatives would need to be listed because the criteria depend on instrument parameters; frankly I'm not yet sure which choices work best in general; and, anyway, this isn't a reference work. One very simple example is worth noting, though. Early in our Treasury Project work, before being fully aware of all the above considerations, we chose the following coefficients for STIS/CCD data:  $A_0 = +0.77$ ,  $A_1 = +0.165$ ,  $A_2 = -0.05$ ,  $B_0 = +0.58$ , and  $B_1 = -0.08$ . These were applied to real pixel values, not function samples  $f(x_i)$ . Fig. 3 shows results compared to linear and spline interpolation used in most software. For any method, the effective output p.s.f. depends on the point source's precise subpixel location as explained earlier. The solid curve in the figure represents a pre-detector p.s.f.; the other curves in the figure are broader and lower because they represent averages over realistic pixels. Each shaded area is an envelope of profiles that occur for all possible subpixel locations.

As Fig. 3 demonstrates, our method (with the coefficients noted above) produces a substantially tighter envelope than either linear or spline interpolation. On the other hand it is not as good as we would like. This figure portrays the behavior of blue-wavelength STIS data quite well, and here I've neglected some details – such as wriggles in the profile wings – too complicated to discuss now. Increasing the number of coefficients produces a modest, not dramatic, improvement in performance.

For more information see <http://etacar.umn.edu>.

## 5. Acknowledgments

The Hubble Treasury Project for  $\eta$  Car is of course supported by funding from STScI. Some of the work reported here was done by John Martin at Minnesota and Kazunori Ishibashi at MIT, with considerable assistance from Matt Gray, Michael Koppelman, J.T. Olds, and R.M. Humphreys. Martin, Ishibashi, and I are grateful to Paul Goudfrooij, Linda Dressel, and Paul Barrett for valuable and interesting discussions of these topics.

## References

- Davidson, K., & Humphreys, R.M. 1997, ARA&A, 35, 1  
 Gull, T.R., Johansson, S., & Davidson, K. (eds.) 2001, Eta Carinae and Other Mysterious Stars, ASP Conf. Ser. 242  
 Humphreys, R.M. & Stanek, K.Z. (eds.) 2005, The Fate of the Most Massive Stars, ASP Conf. Ser. 332  
 Morse, J.A., Humphreys, R.M., & Daminieli, A. (eds.) 1999, Eta Carinae at the Millennium, ASP Conf. Ser. 179