

## Ironing Out the Wrinkles in STIS

Thomas R. Ayres

*Center for Astrophysics and Space Astronomy  
University of Colorado, Boulder, CO 80309*

### Abstract.

The echelle wavelength scales of Space Telescope Imaging Spectrograph were evaluated based on a novel high-density line list for the on-board Pt/Cr–Ne emission lamps. The new reference wavelengths were obtained by a bootstrapping technique that exploited the space-borne instrument as its own “laboratory spectrometer.” A number of strategies were explored to mitigate subtle wavelength scale deviations identified in the process (known from earlier work), either by modifying the pipeline dispersion relations directly, or by a post-facto distortion correction. The main conclusion is that the STIS echelle wavelengths can be significantly improved with only modest changes to the current dispersion model.

## 1. OVERVIEW

Since its installation in *Hubble* more than a decade ago, Space Telescope Imaging Spectrograph (STIS: Woodgate et al. 1998) has been the premier high-resolution ultraviolet spectrometer in space. As such, it has conducted a wide diversity of observing programs on planets, stars, gaseous nebulae, interstellar matter, and even extragalactic sightlines. STIS is especially capable of measuring accurate radial velocities of narrow emission and/or absorption lines, and recording often subtle profile distortions in the spectra of astrophysical objects shaped by a variety of kinematic processes.

Because of its high resolving power, however, STIS is crucially reliant on a robust wavelength calibration, so that the high-precision velocity measurements attainable in principle, can be achieved in practice. For that purpose, the instrument carries a set of wavelength calibration lamps: Pt/Cr–Ne hollow-cathode discharge sources that emit a rich spectrum of very sharp lines in the 1150–3200 Å range where the STIS echelles operate.

The “Wrinkles Project” (aka, “Ironing out the Wrinkles” – Cycle 17 AR–11743) was designed to test the performance of the STIS wavelength scales and identify improvements if warranted. Wrinkles is a follow-on from a Cycle 13 archival program “The Deep Lamp Project,” which examined STIS echelle calibration exposures (“wavecal”) to assess internal precision and absolute accuracy of the pipeline-assigned wavelength scales (Ayres 2008). The method in both projects was to process long-duration exposures of the on-board lamps as if they were ordinary science data; measure positions of high-S/N lines in the calibrated flux density tracings; and compare the recorded values with laboratory wavelengths. The earlier study identified systematic distortions — sometimes subtle, sometimes not — in many of the supported grating ‘tilts.’ An approach had been developed (for the Cycle 14 Legacy Archival project “StarCAT”<sup>1</sup>; Ayres 2010) to correct the distortions post facto, but a central goal of Wrinkles was to evaluate whether an upgraded dispersion model could be implemented in the pipeline itself.

---

<sup>1</sup>See: <http://archive.stsci.edu/prepds/starcat/>

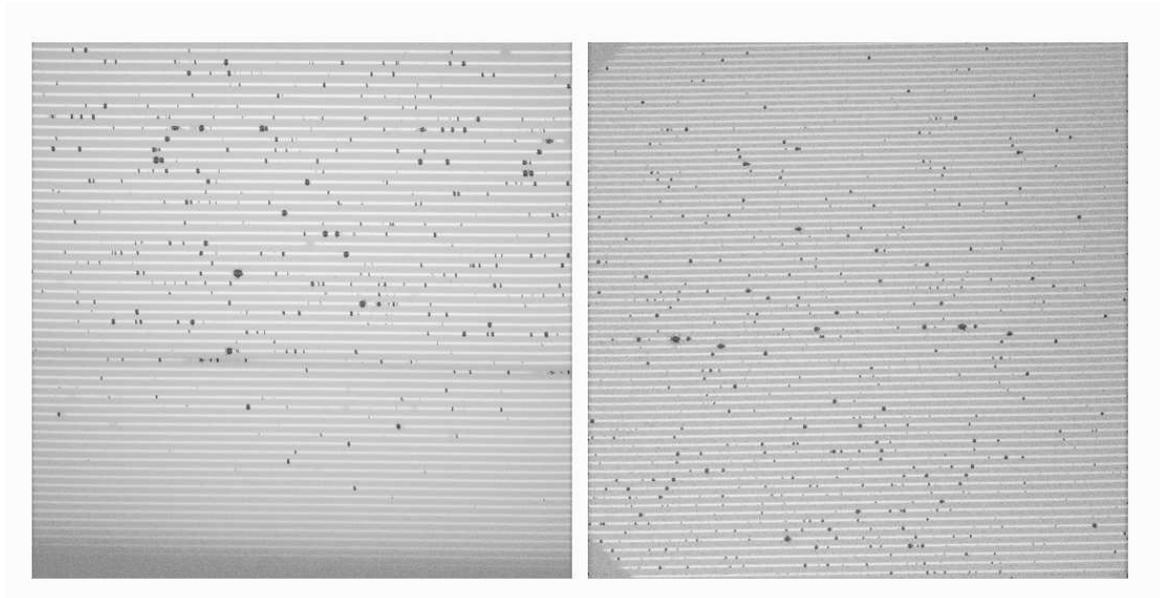


Figure 1: Raw wavecal exposures: E140H-1234 (*left*); E230H-1763 (*right*).

An important advance proposed for Wrinkles was to exploit new laboratory line lists derived specially for the STIS lamps by a partnership between the U.S. National Institutes of Standards and Technology (NIST) and the Space Telescope European Coordinating Facility’s STIS Calibration Enhancement group (see Kerber et al. 2006). This was viewed as critically important for the NUV<sup>2</sup> band, where a rich spectrum of Cr I and Cr II lines appears<sup>3</sup>, in addition to the first and second spectra of Pt and Ne in traditional “platinum lamps,” but could not be utilized directly owing to lack of supporting laboratory wavelengths. Unfortunately, the ST-ECF effort appears to be in stasis as of this writing.

## 2. Laboratory Self-Calibration

Because of lack of progress on the ST-ECF side, the Wrinkles project had to compensate for the pivotal missing Cr I and Cr II calibration material by deriving its own “laboratory” spectrum. This was accomplished as an integral part of the overall analysis, as follows.

First, all the available STIS wavecal exposures longer than 60 s in duration were de-archived (including those taken for re-commissioning STIS following its successful repair during Servicing Mission 4 in May 2009). Examples of representative raw frames for E140H ( $\lambda$ 1234) and E230H ( $\lambda$ 1763) are illustrated in Figure 1. These tilts are at the extreme short ends of the FUV and NUV bands, respectively. Each echellegram covers  $\sim 300$  Å in  $\sim 40$  partially overlapping “orders.” The lamp lines (dark spots) were superimposed on a WD spectrum (light stripes) to show the orders more clearly. The shortwavelength ends of each frame (bottom for E140H, top for E230H) can be recognized by increased crowding of the orders. In both cases, the exposures are not deep enough to capture many lamp features in these high orders, so the dispersion solutions would be less well constrained there.

---

<sup>2</sup>STIS has two general working wavelength ranges for echelle spectroscopy: FUV (1150–1720 Å) and NUV (1650–3200 Å), corresponding to the two MAMA cameras in the instrument.

<sup>3</sup>Chromium was introduced into the STIS lamps to increase the calibration line density at longer wavelengths.

There were 320 datasets in all, counting subexposures: 61 for the single E140M tilt, 69 distributed among the eleven E140H settings, 68 for the six E230Ms, and 122 for the twenty-six E230Hs. About a quarter of these proved to be unsuitable for the wavelength analysis because they had been taken through inappropriate apertures<sup>4</sup>.

Figure 2 depicts the available exposure depth for each setting, counting only wavecalcs taken through the narrow “spectroscopic” apertures. The total cumulative exposure is 37 hours. Each block corresponds to an individual mode/tilt: the width indicates the wavelength coverage, and the height is the relative exposure depth. Dark shaded, color-coded blocks in each of the four sections refer to the prime settings. Thinner, yellow-shaded tilts are ones judged to have poor depth (and candidates for new calibration exposures).

Next, a slightly modified version of the original Deep Lamp pipeline was used to process the datasets, including the cases of multiple subexposures (avoided for simplicity in original Deep Lamp). This involved rewriting the FITS headers of the `o*.raw` wavecal files to reset specific processing keywords so that `calstis` treated the lamp exposure as if it were a normal science image, but taking care to deactivate certain options, like Doppler compensation for the spacecraft motion, which normally are invoked for observations of external targets. Fittingly, the “WAVELINE” (`o*.wav` file, a short [5–10 s] exposure of the lamp preceding a science exposure to set the wavelength zero-point) is a copy of the wavecal itself (first subexposure in case of multiple subexposures). The general procedure has been described by P. Hodge and colleagues (1998).

In the `calstis` processing, a new cross-correlation template initially was implemented, based on the preliminary full-coverage H-mode lamp “atlas” constructed as a byproduct of the Deep Lamp project. In tests, described below, only minor differences were found compared with the historical template used in `calstis`, although the default version does exhibit conspicuous defects in a number of places, as illustrated later (Fig. 4).

The post-processing scheme developed for StarCAT then was applied, in various guises specialized to subsequent analysis steps. The simplest level retained the full 2D order-resolved format of the `x1d` file<sup>5</sup>, either with or without the StarCAT wavelength distortion correction (as described by Ayres 2010). In both cases, the `x1d` spectra were re-sampled to twice the spectral density (to now four points per resolution element [resol]), and the photometric error curve,  $\sigma_\lambda$ , was modified slightly to agree better with Poisson expectations at low counts (as described by Ayres 2010). The uncorrected 2D spectral datasets were the raw material for a new distortion correction, while the corrected echellegrams were used to judge the success of the distortion compensation over the full 2D format.

The second level of post-processing involved merging the orders of the distortion-corrected `x1d` pipeline file into a coherent 1D spectrum for the particular setting, along the lines described in the StarCAT study. The order-overlap zones benefited directly in higher S/N, but more subtly in averaging over presumably uncorrelated residual wavelength distortions, and fixed pattern noise, at opposite sides of the echellegram.

Subsequent steps in the StarCAT protocol involved, first, co-adding subexposures, registered by cross-correlation; then, combining like exposures taken in different epochs. This scheme was applied to the 2D and 1D spectral collections alike. A final processing step, solely for the 1D traces, spliced together all of the individual H-mode wavelength segments (or the M-modes) to yield a coherent “atlas” of the calibration lamp over the full 1150–3150 Å range.

---

<sup>4</sup>The default slits are 0.20×0.09 (in arcseconds) for H settings and 0.20×0.06 for M tilts, although the nonstandard 0.10×0.09 aperture often was used for E230H to mitigate order crowding. In several cases, particularly the 3 prime M settings, the 0.20×0.20 “photometric” slot was utilized. These exposures tend to be brighter, owing to  $\sim 3\times$  higher throughput compared with the default “spectroscopic” slit, but the line profiles also are broader and more “U-shaped” than their sharp Gaussian cousins from the narrow apertures.

<sup>5</sup>Collection of traced spectra for each echelle order of a setting, the main output product of `calstis`.

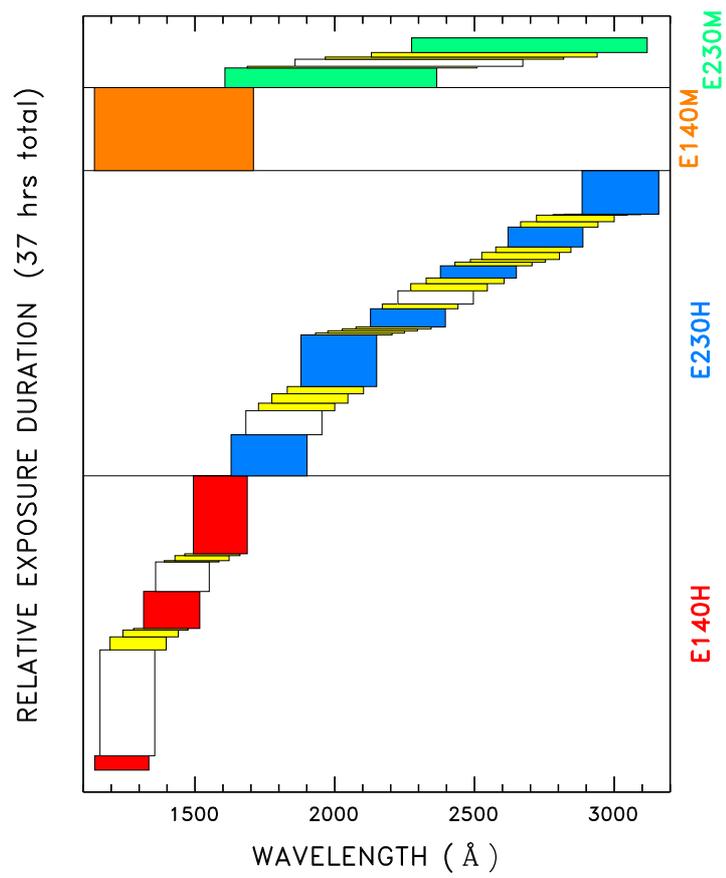


Figure 2: Wavecal exposure depth in each supported STIS echelle mode/setting.

When combining like spectra from different epochs, the wavelength registration step was skipped, because there was no reason to prefer one epoch over the others. In the initial subexposure co-addition, the registration *was* performed, because the reference “WAVE-LINE” for each subexposure defaulted to the leading one of the set, which then would not account for any thermal drifts of the subsequent ones. The post facto cross-correlation procedure takes care of this. (However, any exposure that had an anomalously large shift compared with the majority of a group was excluded from consideration.) Further, in the splicing procedure that combined, for example, the 37 individual high-resolution echelle settings (11 FUV E140H, 26 NUV E230H), there also was no cross-segment wavelength registration (unlike in StarCAT), to avoid biasing the final spectrum by tagging it to positions of a few specific lines across the full range. At most wavelengths the H-spectrum is a combination of several independent echelle settings, recorded on different sets of pixels at different  $y$  positions on the camera, and each individually corrected for wavelength distortions. Because of this, the final spliced spectrum in principle should have higher wavelength precision than any of the individual segments.

Another important departure from the StarCAT protocols involved the treatment of spectral flux densities at the different layers of co-addition. In StarCAT, the calibrated  $f_\lambda$  were weighted in a average by  $\sigma_\lambda^{-2}$  which can be shown to be equivalent to weighting by net counts (see Ayres 2010). This assumes constant intrinsic flux densities (i.e., no variability), regardless of exposure time. For the lamp spectra, however, the intrinsic fluxes are not independent of exposure time, epoch, or aperture; the former two because the lamp output is not constant (changing as the discharges ‘warm up’), and the last because the lamp is a diffuse source. To counteract any bias, the flux densities,  $f_\lambda$ , and photometric uncertainties,  $\sigma_\lambda$ , of a group of lamp exposures selected for co-addition were multiplied by a scale factor to bring into agreement a specific bright line in common to the exposures. The scale factor was set relative to the observation having the highest S/N at the reference feature. A consequence of the procedure is that the independent spectral segments might be offset in flux density from one another, to varying degrees. These offsets were cleared, however, in the final splicing step according to a bootstrap fluxing procedure developed for StarCAT. The flux bootstrapping worked very well, especially for the highly overlapping H settings, because there always were sufficient bright lines in the overlap zones to determine an accurate flux density ratio between adjacent segments.

Figure 3 illustrates a page from an atlas of the resulting H-mode spectrum. The tracings were slightly Gaussian smoothed for display purposes. Note the highly logarithmic ordinate scale. This is required to simultaneously show the strongest lamp lines, together with those at the horizon of detectability, in the face of the remarkable dynamic range achieved in these co-added spectra. Red vertical ticks mark the measured features. Annotations in black are from a hybrid line list derived from NIST laboratory measurements; red are Cr I lines from Wallace & Hinkle (2009); and blue are preliminary identifications based on Atomic Line List v2.04 (more detailed descriptions of the atomic data sources are provided later). The green curve is a heavily smoothed  $5\sigma$  photometric noise level. Only features with peak S/N at or exceeding that threshold were measured in the H tilts.

Figures 4 compares the original calstis cross-correlation template (red curve) with one derived from the new H-spectrum (dots), illustrating a number of conspicuous flaws in the former (note lines missing in the old template, e.g.,  $\lambda 1426$ ). To be sure, this is the worst example of defects in the pipeline template; for the most part, the agreement between new and old is good. When constructing the new template, the H-spectrum was interpolated onto the wavelength scale of the original pipeline version, and the flux densities were scaled by the ratio of the total integrated intensity of the original spectrum to that of the new one. Thus, the new template has the same intensity units as the original, but a somewhat different (likely more accurate) distribution over wavelength.

Following spectral processing, every sharp lamp line with peak  $S/N \geq 5$  was measured in the two sets of 2D records (i.e., the 44 individual frames distortion-corrected or not),

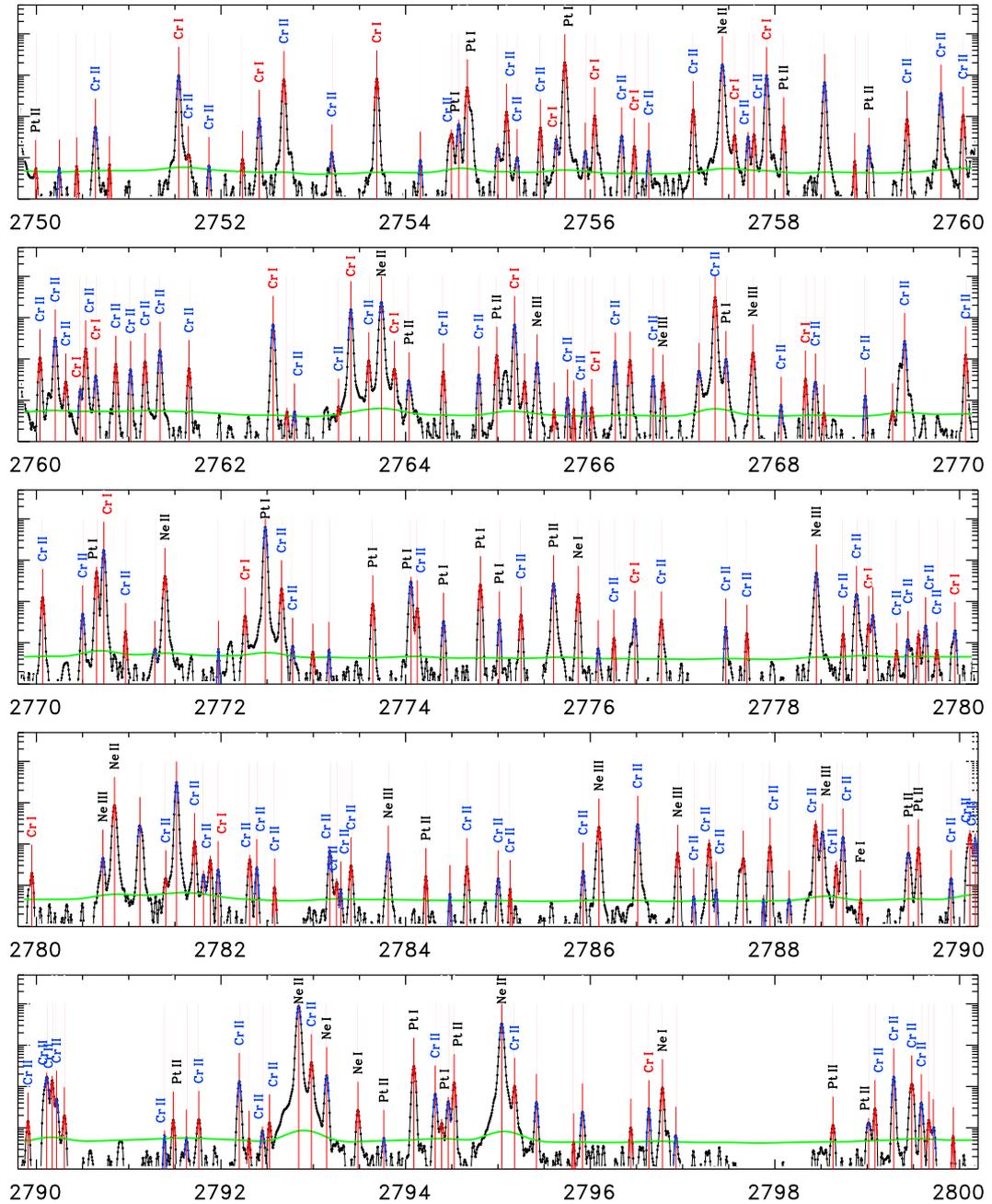


Figure 3: Page from an atlas of highest-resolution version of STIS lamp spectrum.

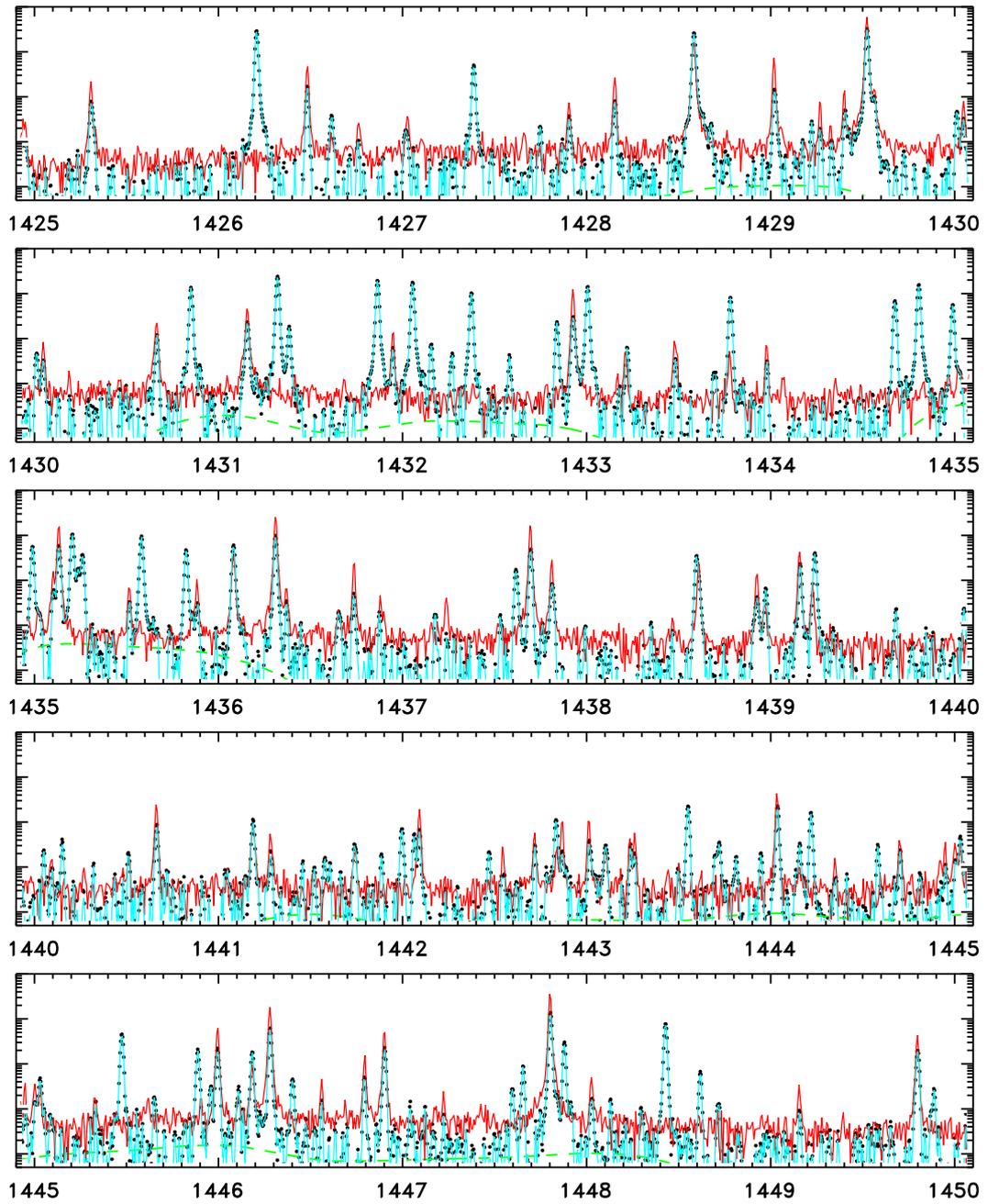


Figure 4: Comparison of new and old cross-correlation templates.

and the single 1D spliced H-spectrum, utilizing an autonomous Gaussian fitting “robot.” The automaton was taught to ignore overly broad features, and weak peaks on the flanks of stronger lines. In general, it was very successful in capturing the truly sharp, isolated emissions prized for calibration, without direct operator intervention (except to ‘tweak’ the governing parameters).

The robo-fitting of the H-spectrum yielded an extensive, refined tabulation of the STIS lamp features. A hybrid laboratory list based on the Sansonetti et al. (2004) measurements of STIS-type lamps (including Cr I and Cr II lines) below 1800 Å, and the earlier NIST study by Reader et al. (1990) of GHRs-style lamps (lacking appreciable chromium) for the longer wavelengths, contains 4335 entries, but some of the features are below the  $S/N \geq 5$  threshold in the co-added Wrinkles lamp atlas, and thus are not useful for defining distortion corrections.<sup>6</sup> The new Wrinkles line positions then were matched to those in common to the hybrid Sansonetti/Reader list (lacking Cr measurements above 1800 Å) and a more recent study of specifically Cr I by Wallace & Hinkle (2009) utilizing the high-precision Kitt Peak Fourier transform spectrometer (FTS), although their wavelengths were restricted to the mid- and long-NUV portion of the STIS band, above 2360 Å.

In comparison to the laboratory measurements, the Wrinkles atlas lines displayed long range systematic behavior, albeit at a low level (few hundred  $\text{m s}^{-1}$ ). The systematic behavior was assumed to be intrinsic to the Wrinkles material, rather than to the laboratory wavelengths, because there was good internal agreement between the deviations displayed by the NIST-classified lines (Pt and Ne) on the one hand, and the Wallace & Hinkle lines (Cr I) on the other, over the wavelengths in common ( $\lambda > 2360$  Å, as mentioned above). A high-order, low-amplitude polynomial was fitted to the deviations and applied to all the Wrinkles wavelengths to put them on the NIST/W–H scale. The corrected Wrinkles list coincides with the corresponding NIST and W–H features (i.e., the average velocity difference is zero) with a standard deviation of just  $\sim 200 \text{ m s}^{-1}$  ( $1\sigma$ ), about *twice* as good as the quoted uncertainties in the laboratory work (suggesting that the latter are conservative).

The procedure was repeated to obtain reference wavelengths for M-mode features, although instead of using the M-spectrum directly (one E140M plus six E230Ms), the high-precision H-spectrum was utilized, after Gaussian smoothing it to the equivalent M-mode resolution. This ensures that reference wavelengths applied to derive distortion corrections for the M modes have the heritage of the heavily averaged high-resolution spectrum, which itself can be tied directly to the original laboratory measurements recorded at comparably high resolving power. Accordingly, the final M line list was corrected for the subtle long range systematic trend identified in the corresponding H-spectrum list.

The smoothing procedure required separate resolution matching for the FUV and NUV regions, owing to the  $\sim 50\%$  higher M resolution of the former. The break point was taken at 1650 Å. In addition, a bright spike at 1187.5 Å was edited out of the H-spectrum prior to smoothing. This feature is an unflagged intermittent detector hot spot on the FUV MAMA, described in the StarCAT study, which survives averaging because there are only two independent spectra that contribute to the sum at these short wavelengths. It probably is present at other wavelengths in the merged spectra, but less obvious because of more widely spaced orders (bright spot perhaps falls in the inter-order background) and the larger number of independent spectra contributing at each final wavelength.

Beyond the numerous matches with the two fundamental reference laboratory sources, the Wrinkles list contains many entries that are not in the NIST and W–H tabulations, mainly the Cr I and Cr II lines that the ST-ECF effort was intended to measure. Most of

---

<sup>6</sup>The  $S/N \geq 5$  criterion translates to a maximum “photometric” measurement error of about  $500 \text{ m s}^{-1}$  in the line positions in equivalent velocity units. The quoted uncertainties of the lab wavelengths are at the level of  $400 \text{ m s}^{-1}$ , although there is evidence that the true  $1\sigma$  uncertainties are smaller.

the “new” features can be assigned preliminary identifications by reference to Cr line lists from atomic databases such as maintained by the University of Kentucky (Atomic Line List v2.04<sup>7</sup>). These new features, regardless of their exact identities, can be utilized — treating the observed wavelengths as “laboratory” values — together with the ones corresponding to the “original” NIST features (but now using the new H-spectrum wavelengths) to refine the distortion maps.

In essence, STIS became its own laboratory spectrometer; bootstrapping from the existing line lists to obtain a better distortion correction, and from that an improved spliced H-spectrum, and from that new, more accurate measurements of the lamp lines. The process was iterated a few times to ensure optimum results. It worked because: (1) the STIS echelle dispersion relations are very linear (§4 below); and (2) the apparent distortions that survive the pipeline processing are large scale and systematic (as shown later).

Figure 5 illustrates schematically the improvement obtained for the H settings (2D spectra). The top panel represents uncorrected *calstis*, as one would routinely retrieve from the MAST archive. The image depicts the density of points (wavelength deviations,  $\lambda_{\text{obs}} - \lambda_{\text{lab}}$  expressed in equivalent Doppler shift), as binned in wavelength and velocity shift. Solid symbols represent averages over 50 consecutive measurements, to illustrate any systematic behavior. The middle panel is for uncorrected *calstis*, but with the new cross-correlation template described earlier. The significant deviation at 1750 Å in the top panel is muted somewhat. The bottom panel is for distortion-corrected *calstis* with the new template. The deviations now are smaller and more uniform, and likely reflect mainly unavoidable measurement errors on both the STIS and laboratory sides. Note that the H mode resol (2.6 km s<sup>-1</sup>) is comparable to the full extent of the *y* axis in each panel.

### 3. Enhanced Distortion Correction

Based on experience developed during this analysis, it was decided to modify the way in which the distortion maps were constructed and applied. Previously, the maps were based on a linear grating parameter “*k*” ( $k \equiv m \lambda$ ) on the *x* axis (the grating parameter varies, in principle, exclusively in the dispersion [*x*] direction) and a linear order “*m*” scale on the *y* axis. However, the orders in the actual echellegram have a nonlinear *y* spacing, which should be maintained in the correction maps to properly capture any coherent spatial distortions related to the detector geometry. A new mapping was undertaken, replacing *m* with a spatial coordinate  $\tilde{y}$  that reflects the true nonlinear order spacing. The  $\tilde{y}(m)$  function was determined by measuring mean order positions (the orders are slightly tilted) in calibration spectra of the DA white dwarf G191-B2B (which had been recorded in all 44 supported echelle settings, and has a bright, nearly line-free continuum spectrum). Further, both independent variables were treated differentially:  $\Delta k = k - k_0$  (with a range of approximately  $\pm 1000$ ) and  $\Delta \tilde{y} = \tilde{y} - \tilde{y}_0$  (with a range 0–1023 [low-res pixels]), where  $k_0$  is a constant for each mode and  $\tilde{y}_0$  is a constant specific to each setting. These new values were further normalized by a factor of 500 to put the difference variables into ranges of order unity, to promote more stable numerical solutions. (The dependent variable describing the local distortions was equivalent velocity shift,  $v$  [km s<sup>-1</sup>].)

The distortion maps were modeled empirically by fully (bi-)cubic functions (16 terms) for the H settings, and by 5th-order polynomials (36 terms) for the denser (in orders per tilt) M settings. In addition, the 2D polynomial corrections were “damped out” beyond certain limits in  $\Delta k$  and  $\Delta \tilde{y}$ , such that at least ten measurements contributed to the solutions *beyond* the cutoffs in each axis to avoid edge effects (i.e., wild variations of the polynomials beyond where they might be constrained by measurements, but still within the valid wavelengths

---

<sup>7</sup>See: <http://www.pa.uky.edu/~peter/atomic/>

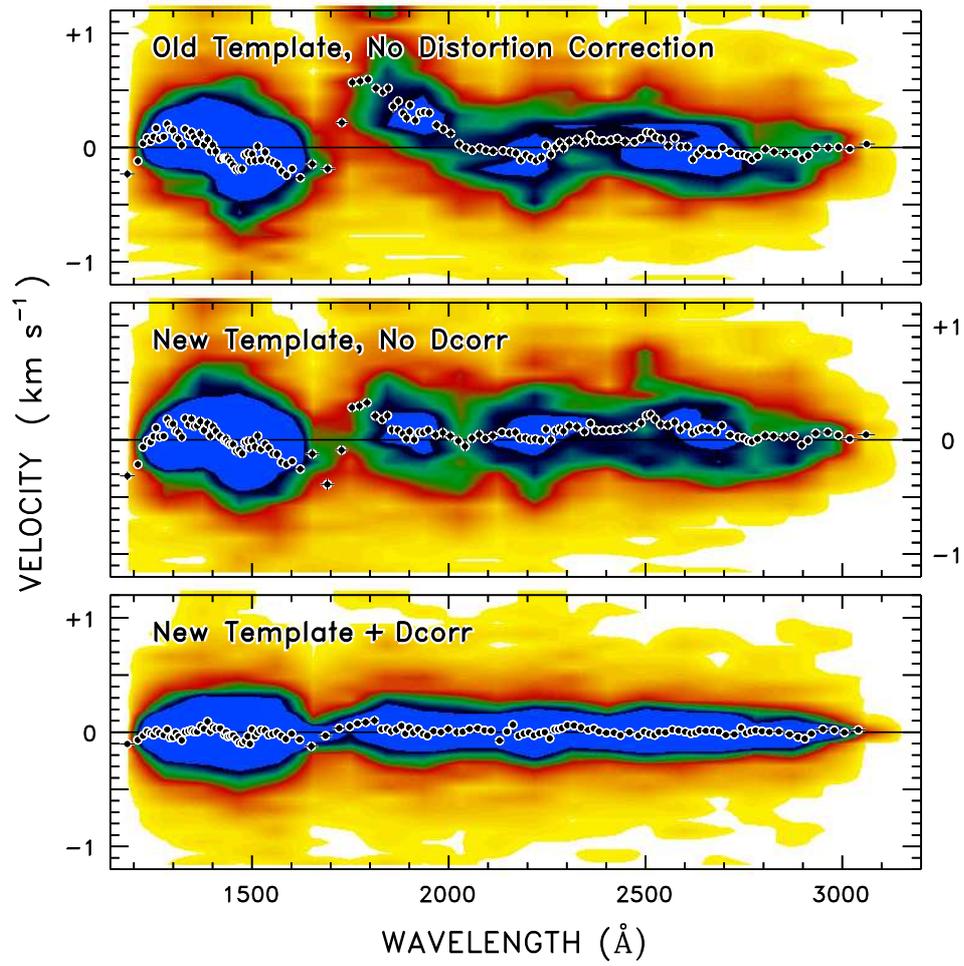


Figure 5: Effects of different processing on 2D order-resolved wavelength deviations.

of the order where the corrections might be applied). Figure 6 illustrates the approach for two representative settings: one M, the other H.

In these figures, the top frame depicts the distribution of measured lamp lines in  $\Delta k$  ( $x$ -axis) and  $\Delta \tilde{y}$  ( $y$ -axis). The size of the symbol is proportional to the absolute velocity shift, and the color codes whether the shift was negative (blue), positive (red), or below a threshold (small open symbols). The velocity cutoffs were  $250 \text{ m s}^{-1}$  for the E140H and E230H tilts,  $500 \text{ m s}^{-1}$  for E140M-1425, and  $750 \text{ m s}^{-1}$  for the E230M settings. These are approximately the limiting velocities that could be measured in principle for a feature of  $S/N \sim 10$ . The maximum velocity shifts typically are only about a third of a resol,  $\sim 1 \text{ km s}^{-1}$  for H settings, and a few  $\text{km s}^{-1}$  for M tilts (see, e.g., Fig. 5 for H). The middle frame visualizes the polynomial distortion model derived from the spatial distribution of velocity offsets in the upper panel, where again blue refers to negative velocities and red to positive. The color scales saturate at  $\pm 4 \times$  the velocity cutoffs mentioned above. The bottom panel illustrates the residual displacements after application of the distortion correction, with the same scaling as the top panel. The orange outlined area indicates boundaries of the region beyond which the models were forced to take on the values along the periphery. Figure 7 summarizes analogous distortion maps for the 44 supported STIS echelle settings. Orientations of the individual frames are the same as in Figs. 1 and 6.

#### 4. Improvements to the Standard Dispersion Model

The next step was to explore whether improvements might be achieved simply by introducing new terms into the existing *calstis* dispersion model, or whether a completely different approach altogether was warranted. As a point of reference, below is the polynomial model incorporated in *calstis*,

$$S = a_0 + a_1 (m \lambda) + a_2 (m \lambda)^2 + a_3 m + a_4 \lambda + a_5 m^2 \lambda + a_6 m \lambda^2 + a_7 (m \lambda)^3 \quad (1)$$

where  $S$  is the “sample” coordinate (echellegram “ $x$ ” axis) in low-resolution pixels,  $m$  is the order number, and  $\lambda$  is the wavelength (in  $\text{\AA}$ ). Only coefficients up to  $a_6$  are populated in the STIS calibration library for echelle settings. Note, also, that the model is not fully bi-quadratic because  $m^2$  and  $\lambda^2$  are missing. Smith (1990) has discussed the venerable *IUE* dispersion relations, upon which the STIS (and GHRs before it) polynomial models were based, including motivation for specific terms from the echelle grating equation.

To carry out dispersion modeling experiments, one needs the distributions of lamp line pixel positions,  $S$ , as a function of order. Conveniently, *calstis* traces the spectrum in raw pixel space, then assigns wavelengths to the pixels by inverting the dispersion relation to infer  $\lambda(m) = f(S, m)$ . Thus, it is simply a matter of re-fitting each lamp spectrum in the native pixel coordinates, rather than assigned wavelengths. The derived line pixel positions should then be in the same coordinate system as the *calstis* dispersion relations, modulo perhaps a constant shift.

As a first experiment, motivated by the Deep Lamp study, the *calstis* model was reformulated using the ostensibly orthogonal independent variables  $k \equiv m \lambda$  and  $m$  (see Smith 1990). The grating parameter  $k$  in principle varies exclusively along the  $x$ -axis, and the order number varies (trivially) exclusively along the echellegram  $y$ -axis, leading to the desired orthogonality. However, a fully bi-quadratic solution (with  $\Delta k = k - k_0$ , where  $k_0$  was taken as constant for each mode) did not significantly improve  $\chi^2$  with respect to a fully bi-quadratic model with terms in  $\lambda$  and  $m$  (which is the *calstis* formulation dropping  $[m\lambda]^3$ , but adding  $m^2$  and  $\lambda^2$ ). The reason for the somewhat disappointing performance will become apparent later.

Figure 8a illustrates the success, or lack of same, of progressively more complex polynomial dispersion models (in  $k, m$ ) applied to the STIS echelle settings. Each shaded box depicts  $\chi^2$  (actually, the “reduced”  $\chi^2$  [per degree of freedom]) of the fit for the particular

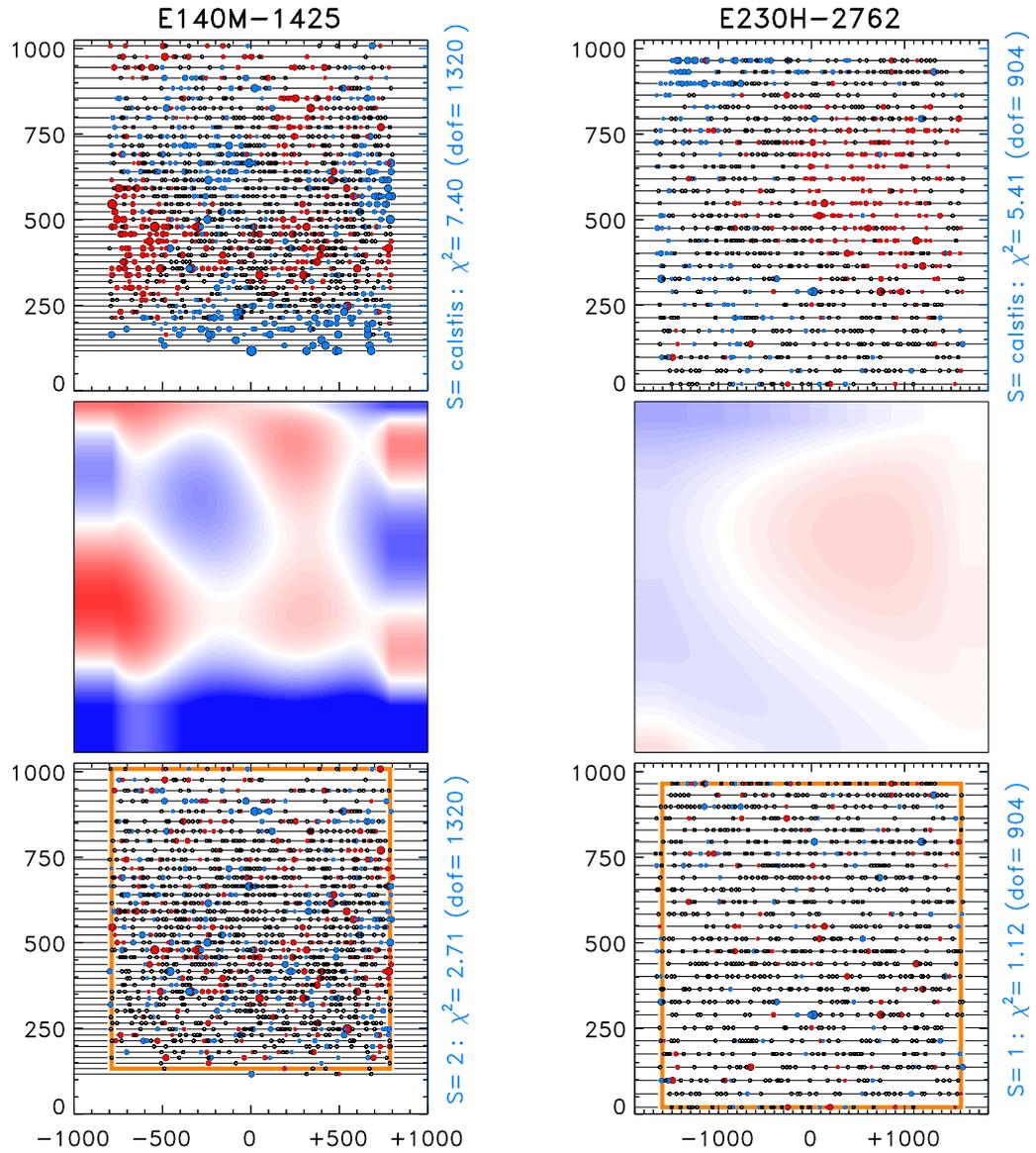


Figure 6: *Left*: 5th-order distortion map, E140M-1425. *Right*: 3rd-order for E230H-2762.

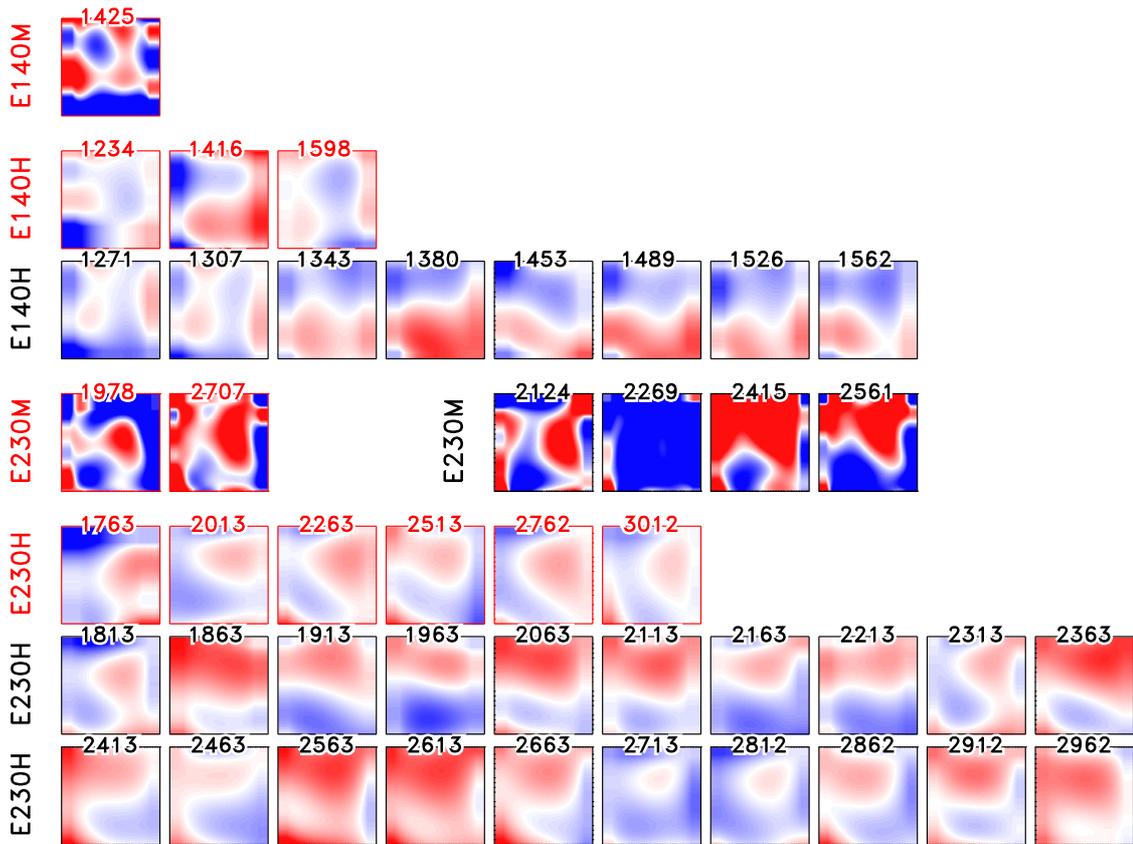


Figure 7: Distortion maps for the 44 supported STIS echelle settings.

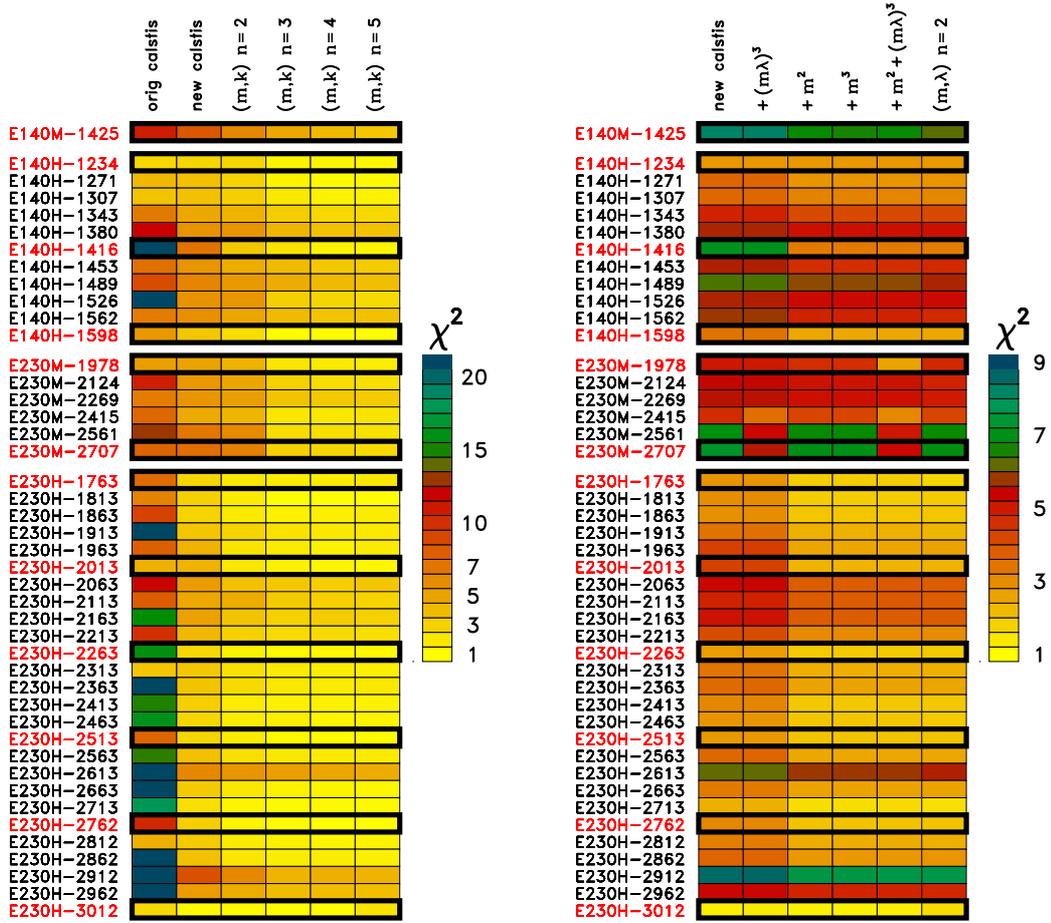


Figure 8: Goodness of fits for various permutations of calstis model.

setting. Prime tilts are indicated in red at left. The lighter the  $\chi^2$  shade, the better the fit. The best examples approach the optimum  $\chi^2 = 1$ . The first column (“orig calstis”) was obtained by taking the normal pipeline dispersion coefficients and calculating the  $S$  values according to  $\lambda$  and  $m$  from the 2D measurements of each setting, then comparing the predicted pixel values with the measured ones (as described above). Before evaluating  $\chi^2$ , the average deviation was subtracted, to account for the fact that the constant coefficient in the pipeline dispersion relations is essentially arbitrary because, in practice, that coefficient is adjusted during the processing to match the zero point indicated by the associated WAVELINE exposure (by the template cross-correlation procedure alluded to earlier).

In general, the  $\chi^2$  for the prime settings appear to be the best (with the possible, although possibly glaring, exception of E140H-1416), but for the secondary tilts generally are worse (and in some cases, very bad). Apparently, coefficients for several of the secondary settings with poor ground calibrations were simply “extrapolated” from neighboring, better characterized prime tilts. In principle, that is a reasonable strategy, but in practice it does not appear to have worked very well. The reasons for this will be described shortly.

The second column (“new calstis”) was the result of re-deriving the dispersion coefficients for the pipeline polynomial, but using the new, more extensive sets of on-orbit measurements now available for the different settings. This provided a fairer baseline for later comparisons than the original calstis coefficients, and illustrates the extent to which that model is capable of reproducing the dispersion properties of each setting, given the spe-

cific (new) wavecal material available. The subsequent columns illustrate the improvement, if any, achieved with higher order polynomials, up to 5th-order. Note that for nearly all the settings, there was not much improvement going from the *calstis* model to fully bi-quadratic (2 additional terms). The 3rd-order polynomial made a noticeable difference for a few of the tilts, mainly the E230Ms. But, aside from E140M, going to 4th- or 5th-order did not improve the  $\chi^2$  appreciably (note that it can even worsen slightly at high  $n$  owing to the decrease in the degrees of freedom tied to the rapidly increasing number of coefficients in the polynomial). The conclusion is that *the relatively simple pipeline dispersion model actually performs pretty well, at least when the coefficients are derived explicitly for each setting.*

Figure 8b illustrates various permutations of the *calstis* polynomial, itself, to test whether modest changes might significantly improve performance. The first column is “new *calstis*” (same as second column in Fig. 8a, but now with a more compact color table). The second column of Fig. 8b turns on the currently unused (at least for the echelle settings) eighth term in the *calstis* model ( $[m\lambda]^3$ ). Except for a few of the E230M tilts, this term apparently is not very potent. The third and fourth columns illustrate what happens when either  $m^2$  or  $m^3$  is added to the pipeline model. Both of these terms appear to be equally improving for many of the settings, especially the numerous E230Hs, but have not much influence on the tilts where  $(m\lambda)^3$  was more effective. The fifth column activates  $(m\lambda)^3$  and  $m^2$  together (equivalent to adding a single new term to *calstis*, if the coefficients for the existing  $[m\lambda]^3$  are populated). This does as well, or in some cases better, than a fully bi-quadratic model (last column). (The bi-quadratic model would add  $m^2$  and  $\lambda^2$  to *calstis*, but take away the  $[m\lambda]^3$  term.) The conclusion here is that *a simple modification to the existing *calstis* dispersion relation could significantly improve performance for all the tilts.*

## 5. Orthogonal Variables and Global Models

There still, however, was the lingering issue of why the supposedly orthogonal independent variable set  $(k,m)$  did not produce the level of improvement over the decidedly non-orthogonal variables  $(\lambda,m)$  anticipated from general principles. The reason became obvious with a simple consideration of the behavior of the dispersion properties of the individual settings. Figure 9 illustrates the representative case of E140H. The left panel displays, order by order, parabolic fits to line centroids ( $S$ : low-resolution pixels) as a function of  $k - k_0$ ,

$$\left[ \frac{S}{500} \right] = a_0 + a_1 \left[ \frac{(k - k_0)}{500} \right] + a_2 \left[ \frac{(k - k_0)}{500} \right]^2, \quad (2)$$

for all orders with ten or more measurements (to ensure an accurate polynomial solution) in all eleven tilts of that mode. If the  $k$  values truly were orthogonal to  $m$ , the curves would fall exactly on top of one another, which clearly they do not.

The lower right panel depicts the coefficients of the quadratic fits. Note that the  $S(k - k_0)$  relation is dominantly linear, with a small but non-negligible parabolic term. Both the linear and parabolic terms appear to be nearly independent of  $m$  in a given setting, and among the settings as well. The constant term, in contrast, displays a clear dependence on  $m$  for each tilt (slanted curves at top of lower right panel).

Another way of looking at this is illustrated in the upper right panel, which depicts the variation of the inferred central  $k_0$  with  $m$  for each setting. (The parameter  $k_0[m]$  is the set of values for a tilt that would force all the curves for that setting in the left panel of Fig. 9 to coincide with one another.) Not only does the “central blaze peak” depend significantly on order number in each tilt, but the whole format apparently shifts systematically (in  $k$  direction) from setting to setting. Here, it appears that  $k_0$  has at least a parabolic dependence on  $m$ , so a bi-quadratic dispersion model in  $k - k_0$  and  $m$ , and with  $k_0 = \text{const}$ , would effectively be missing terms in  $m^3$  and  $m^4$ . At the same time, the apparent dependence of  $k_0$  on  $m$  can be measured, and compensated to restore the proper

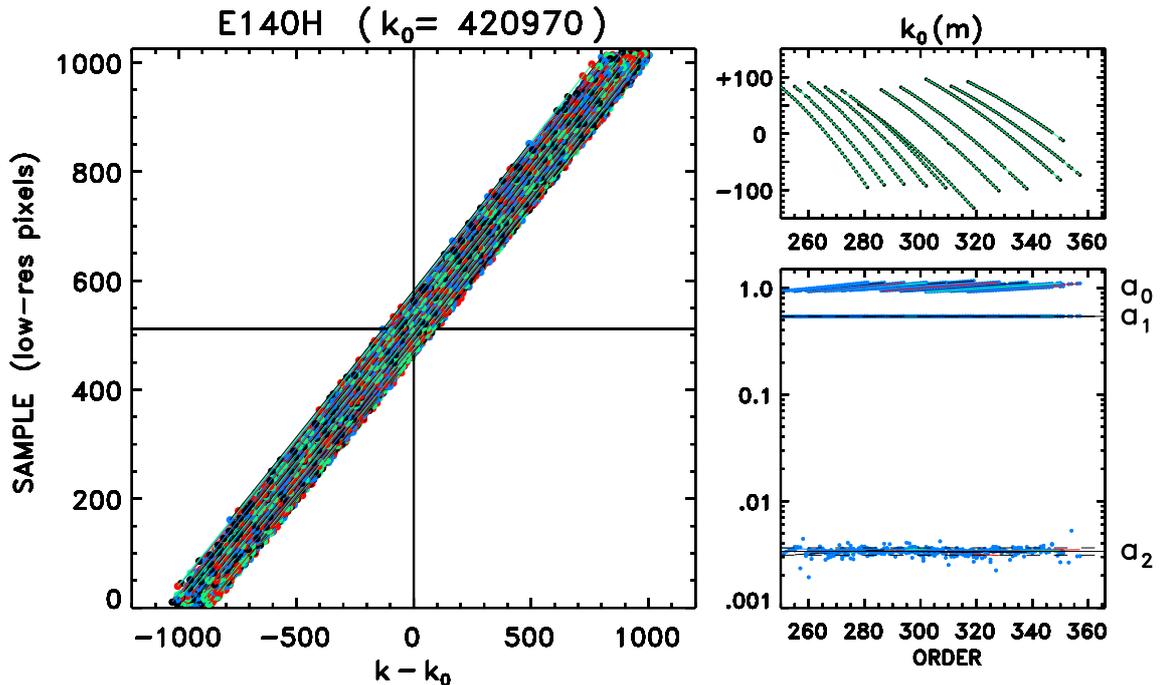


Figure 9: Individual order fits for E140H tilts.

orthogonality of  $k$  and  $m$ . In principle, one could exploit this to solve for a global fit *over all the orders of all the tilts of a mode simultaneously*.

A robust global model would be valuable for the settings currently lacking sufficient exposure depth to achieve a well constrained 2D dispersion solution (yellow shaded in Fig. 2). Without the global model, the next best option simply would be to obtain additional observations in those settings to bring up the exposure depth to a more useful level.

An initial foray into global modeling applied a fully bi-cubic polynomial (16 terms) to all the wavecal lines in all the orders of all the tilts of a given mode (14,000 separate measurements for E230H, for example), but compensating for the apparent  $k_0$  displacements according to relations like those illustrated in the upper right panel of Fig. 9 for E140H. In principle, the  $k_0(m)$  correction should put all the replicas of a given wavecal spot in a given order  $m$  exactly on top of each other in  $S$ , regardless of setting, thereby effectively boosting S/N in a collective fit. It would be as if one could supply a tall MAMA detector that would simultaneously image all the orders of a mode (as E140M does), instead of just a subset. Here, the  $m$  variable was replaced by a spatial variable  $\tilde{y}$  that accounted for the non-linear  $y$  spacing of the orders, and  $\Delta k$ , as in eq. 2 above, again was used, but now taking explicitly  $k_0 = k_0(m)$  as described earlier. This effectively is a “dispersion-only” model, because there is no memory concerning the original location of each measurement in detector coordinates.

Despite the high degree of the model, and promising  $k_0$  correction, the resulting global fits were disappointing. The  $\chi^2$  was 44 for the E230H case, for example. There were large, mixed polarity residuals in diagrams analogous to Fig. 6 (but with  $y$  values extending up to 5700 for E230H, for example, to accommodate all the orders), especially on the extreme left and right edges. The mediocre fits indicated something was missing from the model, probably a systematic behavior tied specifically to the detector reference frame.

A second global experiment attempted to compensate for these putative detector frame effects by introducing six terms to a bi-quadratic form of the dispersion-only model (now,

15 terms total, one short of the fully bi-cubic model used in the initial global experiment). The new terms utilized an independent variable  $\Delta\tilde{y} \equiv \tilde{y} - \tilde{y}_0$ , where  $\tilde{y}_0$  is the  $y$ -position of the first order of a setting at the bottom of the detector. Thus,  $\Delta\tilde{y}$  spans the detector  $y$ -axis (0–1023, in low-res pixels) for each setting, and keeps track of where an order falls physically on the detector for each specific tilt. Again,  $\Delta\tilde{k} = k - k_0(m)$  served as the  $x$ -axis independent variable, but for simplicity  $m$  itself was used for the order-like variable, rather than the  $\tilde{y}$  in the initial global experiment. The final polynomial was:

$$\begin{aligned}
 S = & a_0 + a_1 m + a_2 m^2 + a_3 \Delta\tilde{k} + a_4 \Delta\tilde{k}^2 + a_5 m \Delta\tilde{k} + a_6 m \Delta\tilde{k}^2 + \\
 & a_7 m^2 \Delta\tilde{k} + a_8 m^2 \Delta\tilde{k}^2 + a_9 \Delta\tilde{y} + a_{10} \Delta\tilde{y}^2 + a_{11} \Delta\tilde{y} \Delta\tilde{k} + \\
 & a_{12} \Delta\tilde{y} \Delta\tilde{k}^2 + a_{13} \Delta\tilde{y}^2 \Delta\tilde{k} + a_{14} \Delta\tilde{y}^2 \Delta\tilde{k}^2 .
 \end{aligned} \tag{3}$$

*Astonishingly, this simple approach actually worked.* The  $\chi^2$  (now  $\sim 3$  for previous worst case E230H) was in the range achieved by a bi-quadratic model applied to individual settings having good line densities. The fact that the modeling achieved a dramatic improvement over the initial global fitting efforts (which did not include detector-specific terms) must mean that the large residuals seen earlier do have a very systematic component in the detector frame.

This possibility was investigated by collecting the parabolic fit coefficients for  $S = f(\Delta\tilde{k})$  (the “a” coefficients from eq. 2, but now taking  $k_0 = k_0[m]$ ), averaging them over all the settings of a mode, applying the average transformation to each of the settings individually, and then plotting the residuals in detector coordinates (the  $\Delta\tilde{y}$  mentioned earlier). The results are illustrated in Figure 10.

The apparent quadrupolar distortion pattern is consistent between the M and H modes of each camera, and very similar between FUV and NUV. It is quite distinct from the common types of optical distortions (pincushion and barrel), which would produce a bilaterally symmetric map (redshifts on one side, blueshifts on the other). This probably eliminates the optical path as the culprit and points to the MAMAs themselves. In fact, the inferred distortion maps qualitatively are very similar to that reported in Walsh et al. (2001: their Fig. 11) for the FUV camera, although their companion figure for NUV superficially is quite different. (Note that the Walsh et al. study measured the full vector displacements, whereas here only those parallel to the dispersion [detector  $x$  axis] can be resolved. Also, the Walsh et al. results were for direct imaging with the MAMAs, without the full pass through the complex echelle optics.)

From a practical point of view, introducing an appropriate spatial distortion correction step into calstis would allow the use of a simplified — but potentially more accurate — dispersion model, along the lines sketched earlier, and would permit more reliable “extrapolation” of the model to those settings that have poorer line densities.

## 6. SUMMARY

The central objective of the Wrinkles project was to explore ways to improve precision and accuracy<sup>8</sup> of the STIS echelle mode wavelength scales. Extensive measurements of deep Pt/Cr–Ne lamp spectra in each of the 44 supported modes were utilized to derive 2D distortion maps to correct for large scale systematics traced to limitations of the polynomial dispersion relations embedded in the calstis pipeline. The corrected x1d high-resolution (H-mode) echelle spectra then were merged, averaged, and spliced to produce an improved global atlas of the calibration lamp. This pseudo-laboratory spectrum was measured by an automated fitting ‘robot’ to yield a refined list of the sharp lamp lines, especially the

---

<sup>8</sup>An important component of accuracy, for normal science programs, is the quality of the target centering.

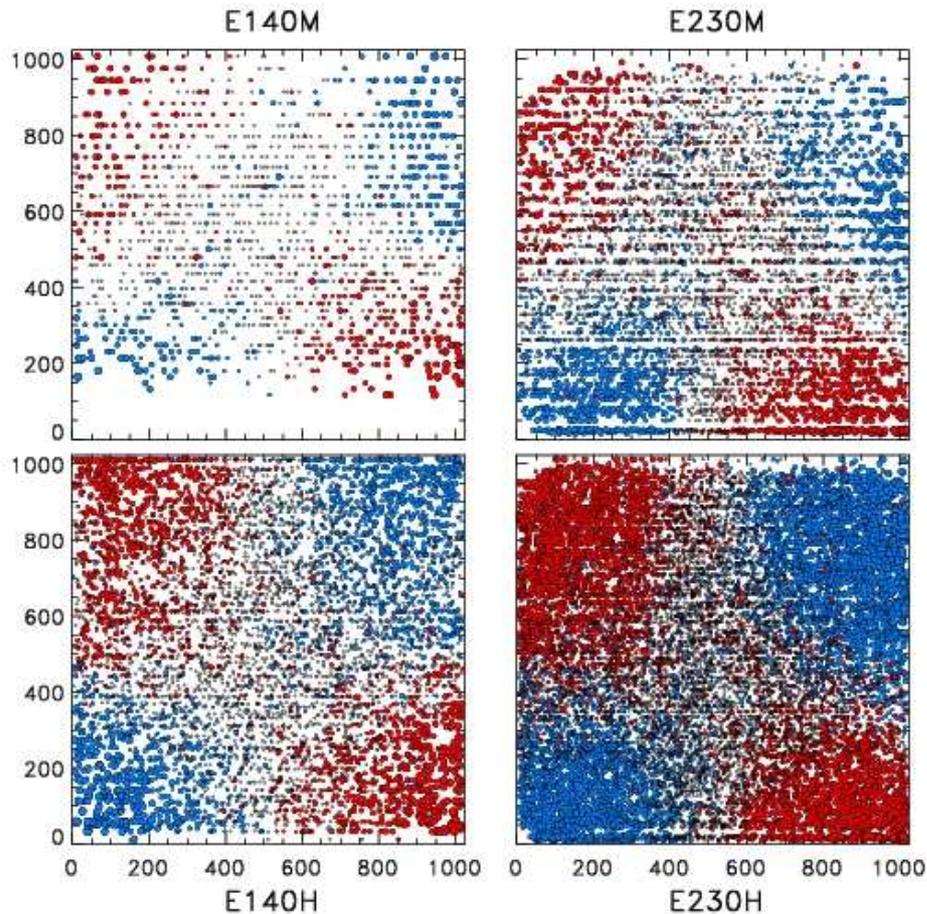


Figure 10: Collective spatial distortion maps for the four STIS echelle modes.

many unidentified features longward of  $1800 \text{ \AA}$  attributed to Cr I or Cr II. The new line list was corrected for low-amplitude, long-range systematics by reference to the subset of features for which accurate laboratory measurements were available. In parallel, the high-quality H-spectrum was substituted for the cross-correlation “template” in the pipeline: the original reference file has a number of conspicuous defects. The high-density line list then was applied to derive second generation distortion maps, and the overall cycle was repeated. Several iterations of the process resulted in a factor of  $\sim 2$  improvement in the (already very good) precision of the STIS echelle wavelength scales, especially for the many secondary settings that are less well calibrated than the primary ones.

A companion analysis considered possible enhancements of the existing pipeline polynomial dispersion relations. The largest improvement was found by simply re-deriving the coefficients for each individual echelle setting, exploiting the now higher density of reliable lamp wavelengths. Further improvement was obtained by adding the term  $m^2$  to render the polynomials almost fully bi-quadratic (only the term  $\lambda^2$  still would be missing), while also activating the vestigial term  $(m\lambda)^3$  (whose coefficients currently are not populated in the calibration library for the echelle settings).

Although it was argued previously that replacing the pipeline independent variables  $\lambda$  and  $m$  with properly orthogonal variables  $k(\equiv m\lambda)$  and  $m$  should improve the solutions, experiments showed, counter-intuitively, that this was not the case. The reason turned out to be that the  $k$  variable, as defined, was not completely orthogonal to the orders

$m$ , because the central value for each order,  $k_0$ , curved slightly along the detector  $y$  axis. Deriving the curvature terms empirically allowed a reformulation in terms of  $k - k_0(m)$ , which not only solved the orthogonality issue, but also allowed the modeling of all the settings simultaneously for a particular mode.

Global bi-parabolic solutions then were obtained, with the simple addition of a few extra terms to account for spatial distortions associated with the detector frame. The quality of the global model was comparable to similar-degree fits to individual settings with the highest line densities. Furthermore, applying average values of the  $k - k_0(m)$  parabolic fits to the individual settings of each mode allowed the spatial (detector frame) distortions to be isolated. The apparent quadrupolar warping is identical in the two cameras, and is similar to independently derived spatial distortions reported for at least the FUV MAMA. If a suitable spatial correction could be introduced into *calstis*, a simplified set of dispersion relations in principle could be utilized.

Third-, fourth-, and fifth-order dispersion models achieved further improvement over the modified *calstis* model. Except for E140M, however, the gains were modest. Indeed, there is some merit to applying the modified *calstis* formulation as an initial wavelength transformation, then following with a subsequent third-, or higher, order distortion correction to account for any low-amplitude residuals. (The latter step might not be necessary if the detector frame spatial correction, mentioned above, could be implemented.)

The main recommendations to the STIS project, in order of increasing importance, if not also increasing difficulty, are:

- (1) Re-derive the *calstis* dispersion coefficients using the updated lamp line lists and the existing on-orbit long duration wavecal exposures. (As of this writing, new coefficients have been derived by the author and are undergoing tests.)
- (2) Improve the wavecal exposure depth for the many modes with relatively poor coverage at present. This includes especially the short FUV end of E140H, and the short NUV end of E230H, where the existing line densities are comparatively low, and the dispersion solutions are less well constrained. (This work is planned for the author's approved Cycle 18 GO calibration program "Deep Lamp Too.")
- (3) Consider extending the pipeline polynomials to a ninth term ( $m^2$ ), and populating the coefficients for the existing eighth term ( $[m\lambda]^3$ ). (The STIS calibration team has agreed to explore implementation of these changes, and suitable coefficients have been supplied by the author for testing purposes.)
- (4) Re-cast the dispersion modeling into two pieces: a de-warping correction followed by a simplified polynomial in orthogonal variables like the  $k - k_0(m)$  approach. With this strategy, the dispersion coefficients for each setting of a mode should be very nearly identical, thereby allowing more legitimate extrapolations to settings with poor line densities. A post facto, high-order distortion correction specific to each tilt could clean up small residuals, if any, from the two-step dispersion correction. (The two-step approach, while the best justified from a "physical instrument" perspective, probably is overkill, given the level of improvement anticipated with step [3], above.)

**Acknowledgments.** This work was supported by archival research grant HST-AR-11743.01-A from the Space Telescope Science Institute, operated for NASA by AURA, Inc. The author thanks the STIS calibration team for numerous helpful discussions. The new lamp line lists and associated spectral catalogs are available from the author upon request.

**References**

Ayres, T. R. 2008, *ApJS*, 177, 626

Ayres, T. R. 2010, *ApJS*, 187, 149

Hodge, P., Baum, S., McGrath, M., Hulbert, S., & Christensen, J. 1998, *HST Instrument Science Report STIS 98–12*

Kerber, F., et al. 2006, *The 2005 HST Calibration Workshop: Hubble After the Transition to Two-Gyro Mode*, eds. A. M. Koekemoer, P. Goudfrooij, and L. L. Dressel, (Baltimore: Space Telescope Science Institute), 318

Reader, J., Acquista, N., Sansonetti, C. J., & Sansonetti, J. E. 1990, *ApJS*, 72, 831

Sansonetti, C. J., Kerber, F., Reader, J., & Rosa, M. R. 2004, *ApJS*, 153, 555

Smith, M. A. 1990, *ESA Special Publication*, 310, 627

Wallace, L., & Hinkle, K. 2009, *ApJ*, 700, 720

Walsh, J. R., Goudfrooij, P., & Malumuth, E. 2001, *Instrument Science Report STIS 2001-02* (Baltimore: Space Telescope Science Institute)

Woodgate, B. E., et al. 1998, *PASP*, 110, 1183