SPACE TELESCOPE SCIENCE INSTITUTE
Operated for NASA by AURA

Instrument Science Report WFC3 2014-17

Updated non-linearity calibration method for WFC3/IR

B. Hilbert
July 18, 2014

ABSTRACT

We report here on the efforts to update and improve the set of coefficients used to correct for signal non-linearity in WFC3/IR data. Using a new method, we have produced multiple sets of correction coefficients, and have compared their efficacy to those coefficients currently used in WFC3/IR calibration. We find that these updated coefficients, derived from a 3rd order fit to the measured linearity versus signal for each pixel, do provide a better non-linearity correction, with photometric results between long and short exposures of the same object more consistent by up to 0.5% compared to the previous correction. In addition, residual non-linearity measured on corrected data is reduced by up to 2% compared to that from the previous correction. Adopting the new linearity calibration will require updates to the WFC3/IR zeropoints in order to maintain the correct flux calibration.

Introduction

HgCdTe detectors, such as that in the WFC3/IR channel, suffer from an intrinsic non-linear response to incident flux. One of the fundamental calibration steps performed by the data reduction pipeline is a correction of this non-linearity. In this report, we detail the method and analysis used to produce an updated non-linearity correction. Previous characterization of the IR channel non-linearity correction are described in Hilbert (2009) and Robberto and Hilbert (2005).

For the work described here, we have implemented a slightly modified method relative to that used in the past. We believe this new method produces a better measure of

the non-linearity, leading to a better calibration result. Details of the method are described in detail in Robberto (2010) and Robberto (2011).

Theory

We used a modified analysis method to create this new non-linearity solution compared to that used previously. Below we give a brief description of the method, and highlight the differences between the old and new methods. The calculations begin with a mean flat field ramp. This ramp is composed of 16 non-destructive readouts of the IR detector. For each pixel, the non-linearity is characterized and appropriate correction coefficients are derived through the use of two iterations of curve fitting to plots of that pixel's signal versus time.

The goal of the first fit is to determine the ideal, linear signal rate that the pixel would measure in the absence of any non-linearity effects. Only by knowing this ideal linear rate can the amount of non-linearity be measured. The differences between the method described in this report and that used in the previous non-linearity analysis lie in the approach used to determine this ideal signal rate.

In the previous analysis, the ideal signal was calculated by fitting a line to only the early, low-signal reads of the ramp, under the assumption that non-linear effects were small at low signal levels and the early reads represented a truly linear measure of the pixel's signal versus time. Once this line was fit, extrapolation using the resulting slope was used to create a set of "perfect" signals versus time for the entire ramp. Details are given in WFC3 ISR 2008-39 (Hilbert, 2009).

For the current analysis, we changed the way in which this "perfect" ramp was created, using instead the method described by Robberto (2010, 2011). This method operates under the more appropriate assumption that non-linearity effects are present in the measured ramp at all signal levels. This implies that fitting a line to the low-signal reads, as was done previously, will result in best-fit signal rates which themselves suffer from non-linearity effects.

In order to account for the omnipresence of the non-linearity effects, rather than a linear fit to the low signal levels we instead fit the polynomial shown in Equation 1 to the full ramp of signal versus time. In the equation below, t represents the exposure time values corresponding to the individual reads, and y represents the fitted signal level up the ramp. At is the product of the exposure time and the constant signal rate that a perfect detector would measure. The C, D, and E coefficients describe the magnitudes of deviations to the set of ideal At signals. These deviations are due to the non-linear response of the detector.

$$y = B + At + C \cdot (At)^2 + D \cdot (At)^3 + E \cdot (At)^4 + \cdots \quad (1$$

For this analysis the polynomial order of the fit was left as a tunable parameter in our script, in order to allow us to explore the quality of the correction derived from various polynomial orders. Details on the fitting process are given in the Analysis section below. Once an appropriate fit is found for a given pixel, the parameter *A* is taken as the slope of the "ideal" linear signal rate, uncontamined by non-linearity effects. This slope is then used to create an "ideal" measure of the signal for each read.

$$y_{ideal} = B + At \quad (2$$

Once the "ideal" ramp signals are calculated, the old and the new non-linearity analysis methods proceed to the second round of curve fitting in almost identical ways. For each pixel, the measured signals up the ramp ($y_{measured}$) are normalized by the calculated "ideal" signals, resulting in a measure of the fractional linearity for each read in the ramp. Note that with this ratio, a measured signal that is unaffected by non-linearity effects would have a value of 1.0. Any non-linearity present in the measured signal would push this ratio below 1.0. Our traditional definition of saturation in the WFC3/IR detector is when this ratio decreases to 0.95.

$$Lin_{measured} = \frac{y_{measured}}{y_{ideal}} \quad (3$$

Note that in the previous non-linearity characterization, detailed in Hilbert (2009), we used 1.0 minus the ratio shown in equation 3. In this new analysis, we found that the fitting and error propagation were easier using the ratio in equation 3 due to the fact that the ratio approached a value of 1, rather than 0, in the case of pixels with very small non-linearity.

We then fit equation 4 to the curve of $Lin_{measured}$ versus $y_{measured}$. As with the previous round of curve fitting, we left the polynomial order as a tunable parameter in our script.

$$Lin_{fit} = 1 - (\alpha + \beta y_{measured} + \gamma {y_{measured}}^2 + \delta {y_{measured}}^3 + \ldots) \quad (4$$

Once a suitable fit is found, the resulting coefficients α through δ can be used to correct the non-linearity for that pixel in any observation, using the equation 5, where s_{meas} is the measured signal in the pixel, and s_{corr} is the signal value corrected for non-linearity.

$$s_{corr} = s_{meas} \times (1 + \alpha + \beta s_{meas} + \gamma {s_{meas}}^2 + \delta {s_{meas}}^3 + \cdots) \quad (5$$

The final correction coefficients for each pixel are saved in a linearity correction reference file, and delivered to Calibration Reference Data System (CRDS, formerly

CDBS) so that the *calwf3* data reduction pipeline can use equation 5 and apply them to all WFC3/IR observations. Details of the calwf3 data reduction pipeline are given in Rajan et al, 2010. By updating this reference file, we are able to update the non-linearity correction applied to all observations.

The current non-linearity correcton coefficients used by the *calwf3* pipeline (which are in the *u1k1727mi_lin.fits* reference file within CRDS) were created from flat field observations made during ground testing. Details of those files and how they were created are given in WFC3 ISR 2008-39 (Hilbert, 2009). These correction coefficients were created from a limited number of ground testing flat field observations. Due to the low signal-to-noise values in the original pixel-by-pixel fits, this current correction uses quadrant-averaged coefficients to correct for non-linearity.

With this updated characterization, we used more data for a higher signal-to-noise correction, and we keep a separate set of coefficients for each pixel, which we believe will provide for a superior non-linearity correction. An important caveat is that, due to the differing methods used to determine the "ideal" signal between the previous analysis and the current analysis, the two methods will measure a given pixel's non-linearity beginning from two different baselines. The differences in the "ideal" signal rate between the two methods will then lead to differences in the amount of non-linearity up-the-ramp, which then propagate into differences in the correction coefficients.

This implies that for a given observation, the application of the new correction coefficients will result in a different set of corrected signals and signal rates compared to what would result when using the old correction coefficients. Using the example of an observation of a standard star, the new correction will produce a different flux for the star than the old correction, meaning a shift in the zeropoint for the IR channel/filter used.

This means that use of the new correction will necessitate an update in the WFC3/IR channel zeropoints in order to give correct absolute photometric measurements. Therefore, prior to implementing any updated linearity correction derived from this work, the WFC3 team will calculate updated zeropoint values. These zeropoints will then be input into the Calibration Reference Database System (CRDS) in conjunction with the new linearity correction, in order to ensure correct photometric results.

Data

In order to create the mean flat field ramps used in this analysis, we used observations collected in the Cycle 18, 19, and 20 IR linearity calibration proposals. See Table 1 in the appendix for details. In all three of these proposals, data were taken in temporally isolated pairs of dark current and internal flat field observations. Internal flat fields were not taken back-to-back in order to minimize any self-induced persistence, which would have complicated the linearity measurements. These three proposals combined to produce 25

flat field observations taken using the SPARS25 sample sequence, plus 3 observations with the SPARS50 sample sequence.

For testing the quality of the new non-linearity corrections, we used observations of 47Tuc acquired as part of the Cycles 17 through 20 linearity monitor proposals. These observations were all made using the F160W filter and had a range of exposure times. See Tables 2 and 3 in the appendix for details.

Analysis

Using the data shown in Table 1, we implemented the new analysis method described above using a series of IDL scripts. The core of the new method, which included the functional form of equation 1, was provided by Massimo Robberto. We then modified that core in order to allow it to function on real on-orbit data. This section provides a details on these WFC3-specific modifications and the final form of the analysis software, followed by a description of the potential non-linearity solutions generated with the software. The software itself is presented in the appendix.

Data Preparation

The data used to generate the new correction coefficients included many flat field observations taken through the F127M filter. With illumination provided by the Tungsten lamp, the mean measured signal rate through this filter was approximately 120 DN/sec/pixel.

Our goal was to measure and characterize the non-linearity in every pixel at signal levels spanning as wide a signal range as possible. The majority of the flat field observations collected for this purpose were created using the SPARS25 sample sequence. After the initial detector reset and zeroth read, this timing pattern samples the signal in the IR detector every 25 seconds. We also included a small number of ramps produced with the SPARS50 sample sequence. This pattern samples the detector every 50 seconds, allowing the 16 reads to monitor the pixels for a longer time. Including this data in the analysis was primarily helpful for those pixels in which non-linearity effects appeared more slowly than average, such that the 352.9 second exposure time of the SPARS25 ramps was not long enough for those pixels to reach saturation.

In order to increase the signal to noise of the observations, we began by creating a mean ramp for each sample sequence. We used the raw version of all observations, meaning that the data had not been run through the calwf3 data reduction pipeline.

With the individual ramps that were used as inputs collected over a time spanning many months, it was important to account for any variations in bias level over time. We did this using the dark current ramp collected contemporaneously with each flat field ramp. For each flat field/dark current pair, we subtracted the zeroth read of the dark current ramp from all reads of the flat field ramp. With the two ramps collected back-to-

back in each pair, the dark current ramp was able to effectively remove the bias signal from the flat field data. Initial tests where we subtracted a single mean zeroth read (created from many dark current frames taken over a long timescale) from all of the flat field ramps showed that contamination from residual bias signal produced less accurate results.

Note that even though we used the dark current ramp as a way to subtract the bias signal from our data, during the polynomial fitting we did not force the *B* term in equation 1 to be zero. This had a small effect on the calculation of $Lin_{measured}$ in equation 3 since the *B* term is present in both the numerator and denominator. The more important difference which can come about due to the value of *B* is when fitting a polynomial to $Lin_{measured}$ versus signal, as the value of *B* will shift the plot in the horizontal direction.

In our case, we did not require *B* to be zero in our fits, as a precaution against any situations where the bias level in the preceding dark current image was significantly different from that of the observation being fitting. In this case, requiring *B* to be zero would have introduced a false non-linearity at the lowest signal levels as the fitted function would have curved down to zero at a time of zero seconds, while the data did not.

Calculating the non-linearity correction coefficients using bias-subtracted data matches with the way in which *calwf3* handles data reduction for IR observations. Specifically, within calwf3 the non-linearity correction is performed after the bias subtraction step. Therefore, it is important for the non-linearity correction to contain coefficients derived from bias-subtracted data. In his work on developing this new method, Robberto derives the non-linearity correction coefficients using data which have not been bias-subtracted, such that the correction includes the signal contributed by the bias, which may make for a more appropriate correction. (Robberto 2011) With this strategy, when applying the non-linearity correction to an observation, it must be done prior to bias subtraction.

After the subtraction of the zeroth read of the local dark current ramp from each flat field, we used all of the flat field ramps taken with a given sample sequence to create a mean flat field ramp. In order to limit the effects of cosmic rays, we used only the data quality arrays present in the calwf3-produced *ima* version of the flat field ramps, where calwf3 had flagged the appearance of each cosmic ray. For each pixel within each ramp, we ignored any reads which containted the 8192 (cosmic ray) flag. This means that for a pixel containing a cosmic ray, we ignored not only the read where the cosmic ray first appeared, but all subsequent reads. This was done in order to avoid any cosmic ray-related effects (persistence, inconsistent signal rate, etc.) that may have been present later in the ramp.

For each pixel in each read, we then calculated the sigma-clipped mean value and standard deviation of the signals in the 28 SPARS25 ramps (or the 4 SPARS50 ramps). These values were saved as the mean signal and associated uncertainty. We repeated the

process for the SPARS50 observations, in order to produce a mean SPARS50 ramp. These mean ramps were then used together as inputs in the new non-linearity analysis method described in the previous section.

Non-Linearity Calculation Details

With the appropriate data files in hand, several preliminary checks were necessary prior to performing the curve-fitting descibed above. First, we needed to limit the signal range over which we performed our curve fits. In general, once the signal level in a given pixel reaches the full well depth, the measured signal in later reads can display odd behavior, increasing or decreasing at different rates unrelated to the actual flux impinging the pixel. These unrealistic measures of the signal had to be removed prior to the polynomial fitting. We therefore needed to identify pixels where full well had been reached and for each pixel, throw out any reads collected after this point, in order to avoid having the curve fits attempt to fit to them.

We employed several filters to find and remove these reads. First, any pixels which had a maximum measured signal that occurred within the first two reads were declared to be saturated for the duration of the ramp, and were completely ignored. (This filter in fact caught pixels showing several odd behaviors, including being saturated all the way up the ramp, or measuring decreasing signal with time. In all cases, these pixels were considered pathological and ignored.)

Similarly, any pixels which never accumulated more than 100 counts during the ramp were declared dead and ignored.

The next filter we implemented was designed to catch the reads later in the ramp where a pixel was approaching or beyond full well. We calculated the mean signal rate in the first 3 reads of the ramp, and then used that rate to estimate the expected signal values in all of the following reads. This is similar to the generation of the "ideal" ramp done in the old version of the non-linearity analysis. In this case we subtracted the measured signals from those in the calculated ramp, and declared that any reads where the measured signal was more than 25% below the calculated value were close enough to full well that they would be ignored. Note that the traditional definition of saturation for the IR channel is the signal level corresponding to a 5% non-linearity, so the 25% cutoff value only removes reads which are far beyond the signal level that will eventually be calculated as the saturation limit for that pixel. Also note that past work (Hilbert, 2009) has shown significant pixel-to-pixel variations in the full well depth, which was the motivation to apply these filters rather than declare a single signal value as saturation and flagging all reads with signals above this value.

We also experimented with ignoring the 0^{th} , or the 0^{th} and 1^{st} reads for each pixel, in order to limit any reset effects present in the early reads of the data. The number of reads to be skipped was left as a tunable parameter in our scripts, in order to use the data to determine the optimal value.

Once these filters had been used to limit the data to good measures of the signal, we performed the first round of curve-fitting. For each pixel, the signal values and accompanying exposure times for the SPARS25 and SPARS50 ramps were combined, and equation 1 was fit to the data. We left the polynomial order to be fit as a tunable parameter (where the user can specify the maximum polynomial order to attempt). In addition, for pixels where the number of reads to be fit was severely limited by the filters described above, the order of the polynomial fit was automatically decreased from the user-supplied maximum in order to allow the fit to proceed.

Once the data to be fit were finalized, and the polynomial order of the fit was known, we proceeded with the fit using the mpfitfun.pro function from Markwardt (2009). This function allowed us to input uncertainties in the measured signals, and to calculate a reduced χ^2 value for each fit. Each time a fit was performed, the resulting χ^2 value was compared to a user-defined upper limit. If the measured χ^2 value was less than the upper limit, then the fit was declared good. If it was above the upper limit, then the final good read (that with the highest exposure time) was chopped off of the inputs, and the fit was attempted again. This process was repeated until a χ^2 value less than the upper limit was achieved. In the case where the χ^2 values of the fits never decreased below the upper limit, we reverted to the fit with the lowest χ^2 value.

The motivation behind this strategy was that the rapidly increasing non-linearity at the highest signal levels was the reason behind the poor fitting results. By removing the highest signal read, we can move away from this rapidly changing regime and give the polynomial a better chance at a successful fit. In theory one could fit higher up into this rapidly changing regime using a higher polynomial order, however we found that when the polynomial order became too large, the quality of the fit at low signal levels began to degrade as the script attmpted to fit to the large changes at the high signal end of the ramp. As the current architecture of the non-linearity correction algorithm within *calwf3* expects a single set of polynomial coefficients for each pixel in order to perform the correction, we are limited to this method for correcting the non-linearity.

In figure 1 we show a typical fitting result. The top panel shows the measured signal values (black diamonds), along with the best-fit result (blue line), in the top panel. The two right-most measured signal values, denoted by the black diamonds without the interior red point, were removed from the fitting due to the high χ^2 value. In the bottom panel, we plot the residuals from the best-fit. Dark blue points show data from the SPARS25 mean ramp, while light blue points are from the SPARS50 mean ramp.

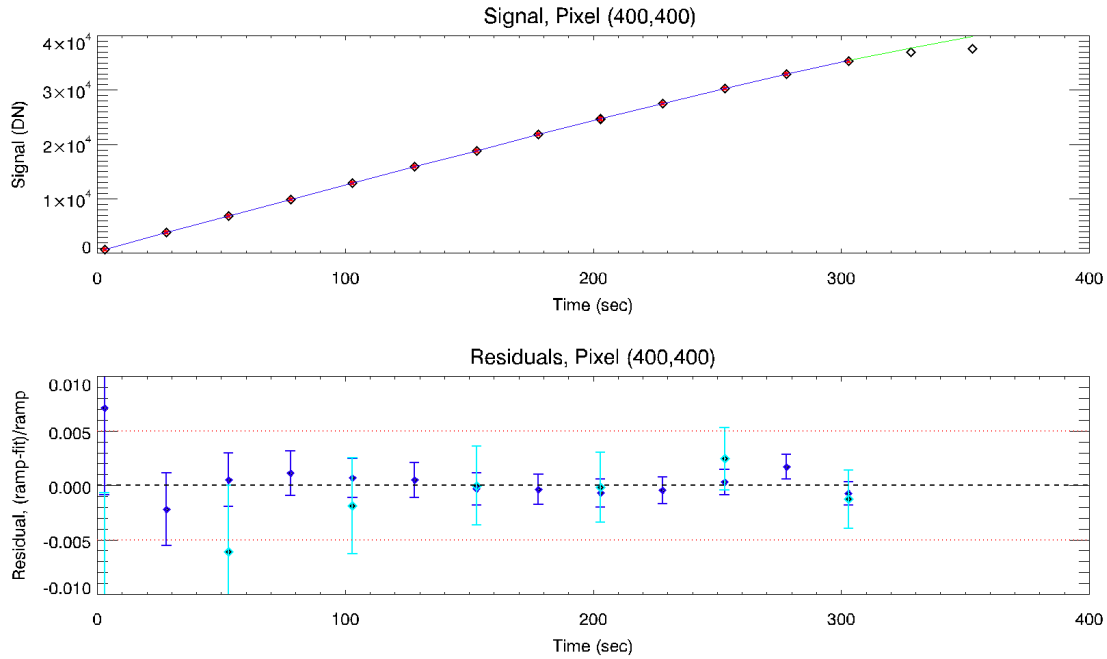


Figure 1: Top panel: best-fit plot to the signal versus time for a single pixel. Only points with the red dots were used in the final best-fit solution. The two higher signal points resulted in fits where the χ^2 value was above the user-set maximum limit. Bottom panel: Residuals from the best fit curve.

Once a good fit was found for a given pixel, we used only the fitted slope and intercept coefficients to generate an "ideal" linear ramp and linearity values, following equations 2 and 3, respectively. The next step was to fit equation 4 to the plot of $Lin_{measured}$ versus $y_{measured}$. We followed the same general strategy for this fitting as that described above, using mpfitfun.pro to fit a polynomial to the data. A user-defined maximum polynomial order and maximum reduced χ^2 value were again used to guide the fitting, and the strategy of iteratively removing the highest-signal read and repeating the fit helped assure a good result. Figure 2 shows an example fit to this data for the same pixel shown in figure 1 above.

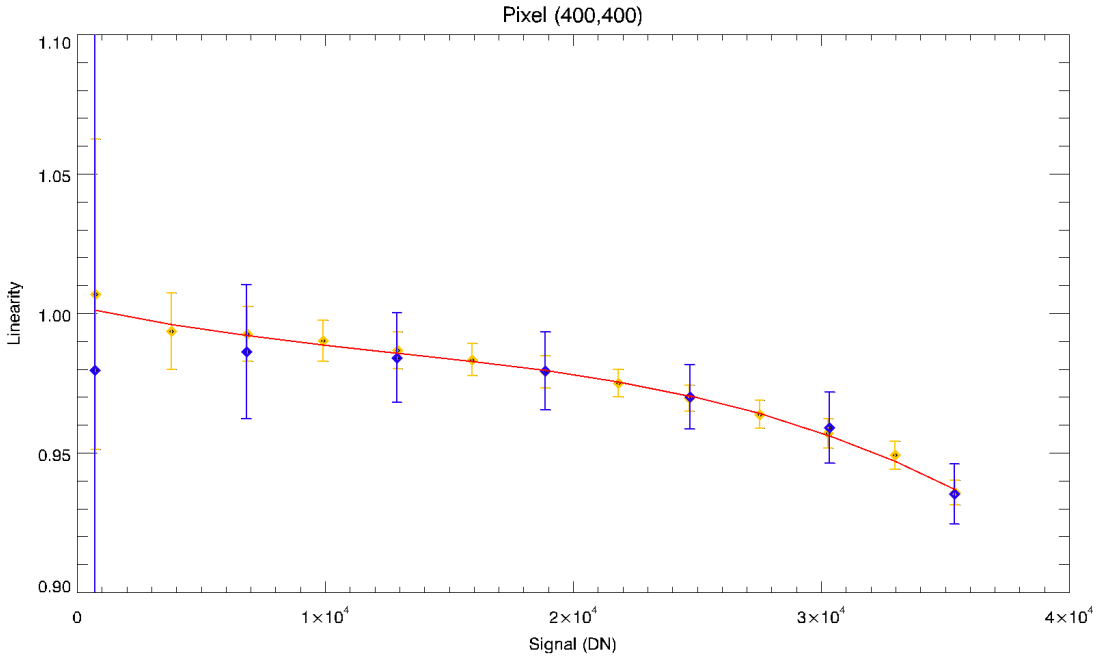


Figure 2: Example plot of the best-fit curve to the linearity versus signal for a single pixel. Blue points represent SPARS50 data, while the yellow points show SPARS25 data. In this case, no high-signal data points had to be removed in order to achieve a good fit.

Using the mean flat field ramps and this method, we then began our search for the best set of WFC3/IR non-linearity correction coefficients. We ran the analysis above many times, while varying the maximum orders of the two polynomial fits, as well as the two cutoff values for χ^2. Each run of the software produced a different set of coefficients. The next piece of our analysis was to determine which set of coefficients resulted in the best non-linearity correction, as measured by the resulting linearity of a set of corrected data.

Comparing Results

In order to determine which set of non-linearity correction coefficients provided the best correction, we applied several new sets of coefficients, as well as the coefficients currently in use in CRDS, to several types of IR observations and measured the linearity of the corrected data. For this analysis, we focused on the three new sets of non-linearity coefficients which appeared most promising.

The first set of coefficients were generated using a 4^{th} order polynomial fit in equation 1, a 3^{rd} order fit in equation 4, and χ^2 limits of 0.04 for both. For brevity, we refer to this set of coefficients as the "430" correction, indicating the orders of the polynomial fits, along with the fact that we had ignored the 0^{th} read when calculating the coefficients.

The second set of coefficients analyzed were produced using a 6^{th} order fit in equation 1, and a 5^{th} order fit in equation 4. These coefficients make up the "650" correction.

Similarly, we refer to the final set of new coefficients investigated as the "670" correction, resulting from 6^{th} and 7^{th} order polynomial fits in equations 1 and 4 respectively. These three sets of polynomial fitting parameters returned the fits with the lowest residuals out of our many trials.

In all three cases, the "0" refers to our finding that skipping only the 0^{th} read of the input ramps tended to produce the best fitting results. In cases where the 0^{th} and 1^{st} reads were both skipped, the lack of data at small exposure times allowed the fitted curves to vary too much at low signal levels, resulting in poor fits.

Comparison Using an Internal Flat Field Ramp

We began with an internal flat field ramp, as this represented data that was very similar to the data used to create the coefficients. We chose to use ib9m08a3_raw.fits, which is a flat field observation taken using the F160W filter in the Cycle 17 Internal Flat Field Monitor program (proposal ID 11915). Note that this is a different filter than that used to obtain the observations we used to create the correction coefficients (F127M). This broader filter resulted in a ramp with a mean signal rate about two times higher than that in the F127M data. Figure 3 shows a histogram of the signal rates in this F160W flat field file.

We ran calwf3 on the file once for each non-linearity correction file, producing a separate *ima* and *flt* file for each set of correction coefficients. As the *ima* files captured the corrected signal levels for all reads up the ramp, our initial analysis focused on these files. We began by examining individual pixels. For a given pixel, we simply plotted the measured signal rates (calculated by dividing the signal differences between consecutive reads by the appropriate exposure time) up the ramp versus time. A ideal pixel that had been perfectly corrected and without noise, would show no change in the signal rate for the duration of the ramp and therefore display a horizontal line.

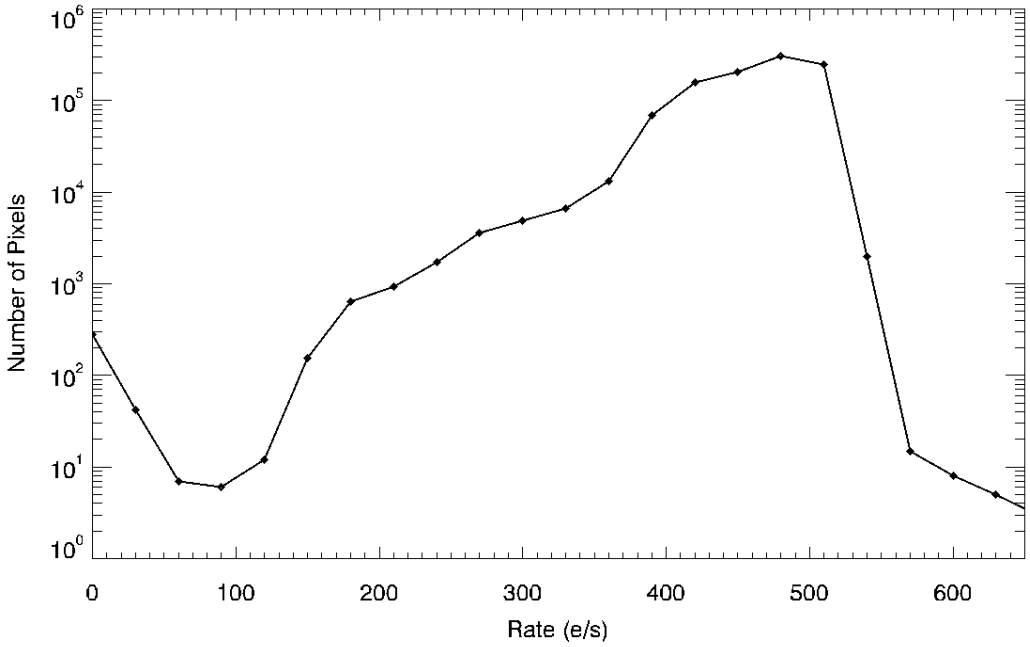


Figure 3: Histogram of signal rates in the F160W flat field file used to test the non-linearity corrections. We binned pixels by signal rate for this analysis, and focused on rates between 150 and 550 e-/s. Outside of this range the limited number of pixels led to noisy results.

By overplotting the rates for this pixel resulting from each of the non-linearity corrections, we can see which correction gave the best result. Figure 4 shows an example of this plot for pixel (962,141), which has a signal rate of about 320 e-/sec. In this plot, the dashed black line shows how quickly the signal rate drops when no non-linearity correction is applied. The solid black line shows what happens to the signal rates when the current CRDS non-linearity correction is applied. This line is much closer to horizontal than the uncorrected line. The red, green, and blue lines on the plot show the signal rates when using the new 430, 650, and 670 coefficients, respectively. All three of these corrections produce a similar result for signal rate up the ramp. Figure 5 shows a similar plot for a pixel with a higher signal rate. In this case all of the corrections appear to leave a small residual slope in the data.

In order to determine which correction gave the best overall results, we needed to come up with a way to measure any residual slopes or curvatures to data such as that seen in figures 4 and 5. For this, we made two fits to the plots. First, we fit a line to each measure of corrected signal rate versus time, and used the devitaion of the fitted slope from zero to guage the quality of the non-linearity correction. In addition we fit a quadratic to a plot of the corrected signal versus time, in an attempt to be more sensitive to residual non-linearity effects present towards the high-signal end of the ramp. We used the coefficient of the quadratic term of the fit as a proxy of residual non-linearity.

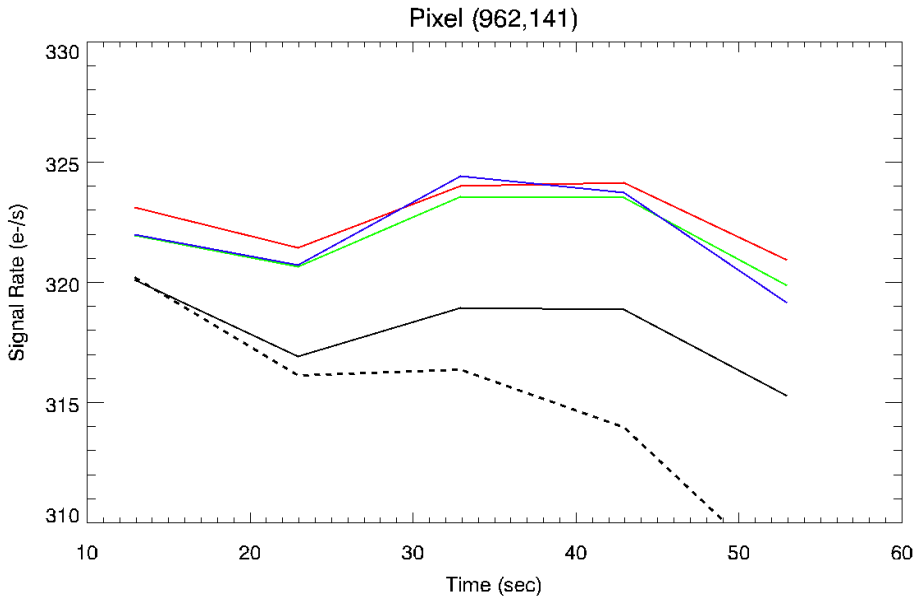


Figure 4: Example correction of a pixel with "mid-range" signal rate. The dashed line represents the signal rate when the ramp was not corrected for non-linearity. The solid black line shows the results of the correction currently used in the CRDS database. The red, green, and blue lines represent potential updated corrections. A perfectly corrected pixel would have a constant rate over the course of the ramp, and be seen as a horizontal line in this plot.

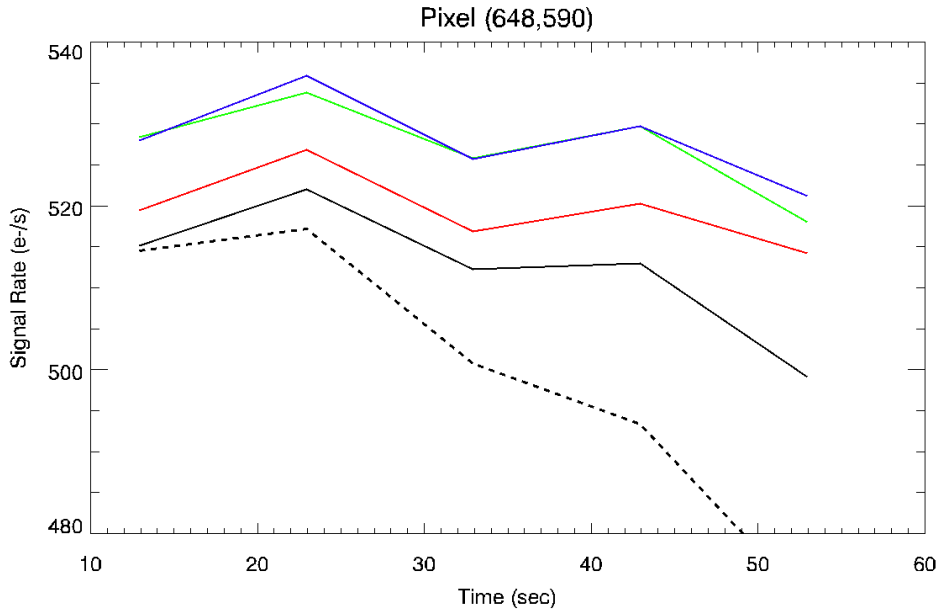


Figure 5: Example corrections on a pixel with a higher signal rate. The color scheme is identical to that in figure 4 above. Note that in this case none of the corrections completely remove the effects of non-linearity, as can be seen by the slight residual negative slope in all the lines.

We made the linear and quadratic fits, and collected slopes and quadratic coefficients for every pixel in each corrected ramp, as well as an uncorrected version of the ramp. In order to interpret these sets of coefficients, we binned the pixels by signal rates up the ramp, and calculated the mean slope or quadratic coefficient for each bin. This helped to improve the signal-to-noise of the measurements, and in the case of the uncorrected ramp, potentially helped to uncover any rate-dependent differences in non-linearity behavior.

Since this first analysis is being done on a flat field, there was a limited range of signal rates represented in the data. Based on the histogram in figure 3, we chose to focus our analysis on pixels with rates between 150 and 550 e-/s, as this range contained the majority of the pixels and allowed for many pixels in each signal rate bin.

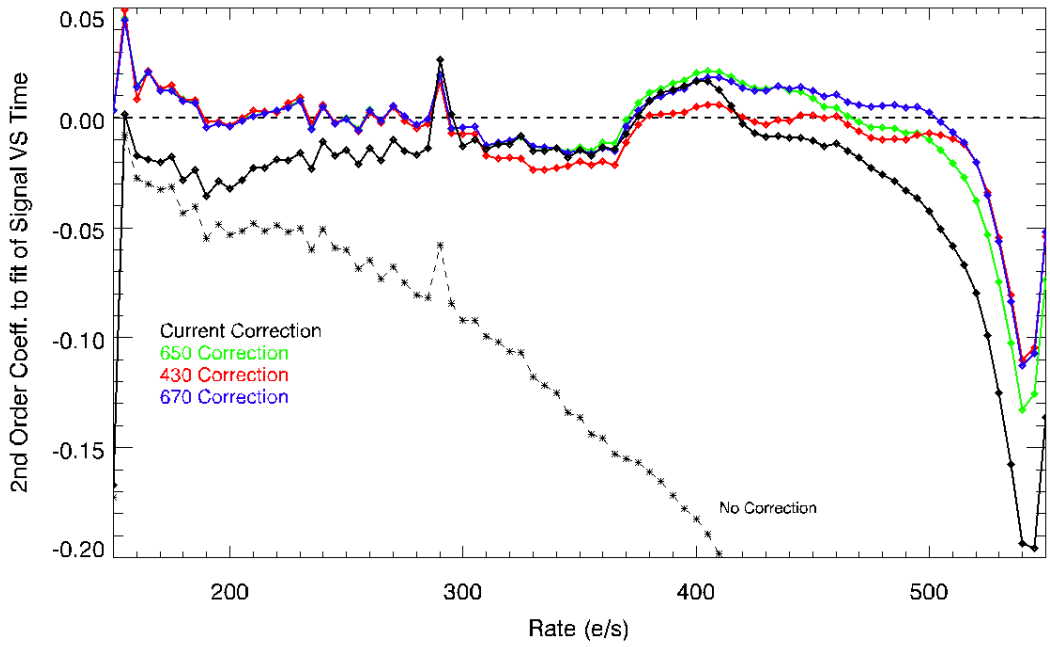


Figure 6: Quadratic coefficients from fitting 2nd order equations to the signal versus time in the corrected files. Perfectly corrected pixels would have a value of zero in this plot. All corrections greatly improve the linearity of the signal compared to applying no correction (dashed line). At almost all signal rates, the new corrections give superior results to the current correction.

Figure 6 shows the mean measured quadratic coefficients plotted versus the signal rate for the flat field file. Again, the dashed black line shows the results for the uncorrected ramp, where the full effect of the non-linearity can be seen. In this case, the quadratic coefficient decreases quickly as the signal rate increases. This is as expected, since pixels with higher signal rates will increase their measured signal more rapidly and begin to exhibit non-linearity effects more quickly than pixels with lower signal rates.

With non-linear effects becoming large earlier in the ramp, these pixels will have much more curvature in their plots of signal versus time and will therefore have a larger (absolute value) quadratic coefficient.

The solid back line in figure 6 shows that the current CRDS non-linearity correction removes much of the curvature present in the uncorrected data. However, for signal rates from roughly 50 – 360 e-/sec, the current correction does leave behind a small amount of curvature in the signal versus time plot, with the mean quadratic coefficient remaining at -0.02 to -0.03 for those pixels. The red, green, and blue curves, representing data corrected using the new non-linearity corrections, show quadratic coefficients which are much closer to zero in this signal rate range.

In addition, at signal rates above about 420 e-/sec, the current non-linearity correction begins to fail, with the quadratic coefficient falling quickly as the signal rate increases. With the new corrections, this breakdown does not begin occurring until roughly 500 e-/sec. At these high signal rates, the blue and green curves are often above zero, implying that the 650 and 670 coefficients are in fact over-correcting at these locations. The red line of the 430 correction appears closest to zero for most of this high signal rate regime.

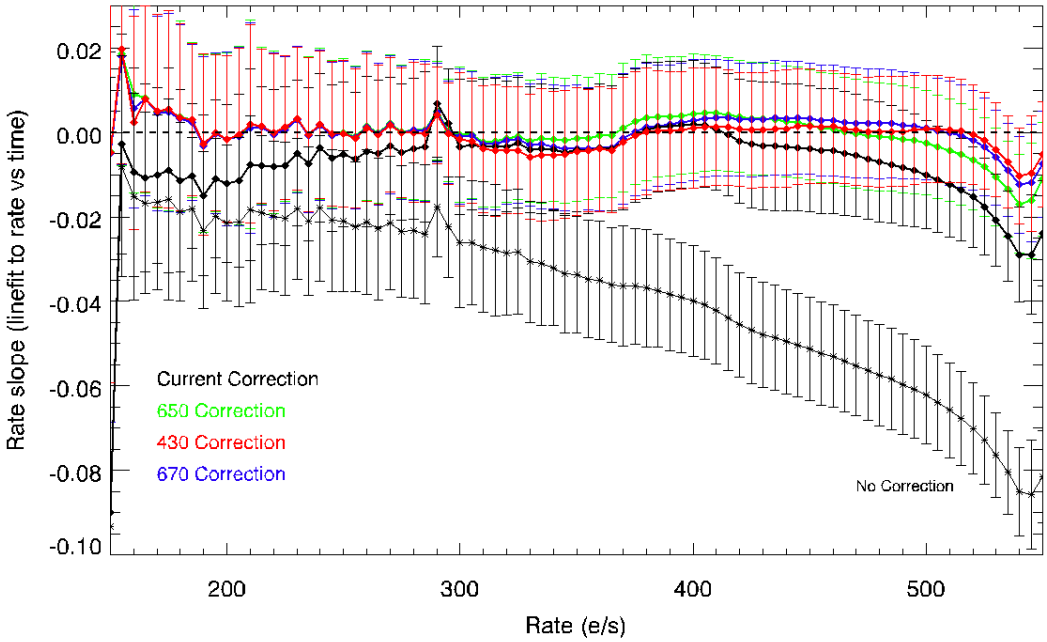


Figure 7: Plot of the slope fitted to the rate versus time up the ramp, versus the observed rate early in the ramp for all pixels. A perfect linearity correction would lead to values of zero. Another way to read the y-axis is to consider it the fractional difference in the rate between the beginning and the end of the ramp. For example, with no correction (lowest curve), the average pixel that begins a ramp with a rate of 500 e-/s will have a rate about 6% lower, or 470 e-/s, at the end of the ramp.

Figure 7 shows the results of fitting a line to the signal rate plots, such as those seen in figures 4 and 5, for the F160W flat field file. The results here are similar to those seen in the quadratic coefficients in Figure 6. At lower signal rates, for pixels less than 285 e-/sec, the current correction (solid black curve) appears to under-correct, while all three new corrections result in slopes much closer to zero. The current correction also appears to under correct at rates above 420 e-/sec, while the 650 and 670 corrections provide a slight over correction. The 430 correction appears to give the best results, keeping the coefficients closest to zero.

Comparison Using External Observations

The next step was to compare the non-linearity correction coefficients when applied to external observations rather than an internal flat field. We performed this comparison on an F160W exposure of Omega Centauri, in order to place many point sources of widely varying brightness across the entire detector. Specifically, we used observation ibcj01ttq_raw.fits, from the Cycle 17 IR Low-frequency flat program (proposal 11928). As with the preceding flat field observation, we ran the Omega Centauri observation through the *calwf3* pipeline 5 times: once using each of the 3 new non-linearity corrections, once using the current correction implemented in CRDS, and once with no correction. With the much wider range of rates present compared to the flat field data, this observation provided a much more complete look at the effects of the various non-linearity corrections. Figure 8 shows a histogram of the rates observed in the early reads of the ramp.

Figure 9 below shows the quadratic coefficient of the fits to the signal versus time in the corrected ramps. The results here are significantly different than those for the F160W flat field ramp, seen in figure 6. While the dashed curve of the uncorrected ramp is fairly similar between the two cases, the shape of the curves for all four corrections are different. In the case of the flat field, the curves associated with all four corrections were relatively flat and close to zero up to signal rates just under 500 e-/sec. But in the Omega Centauri ramp, all of the corrected ramps have quadratic coefficients much farther from zero. For example, for pixels at a signal rate of 450 e-/sec, the coefficients associated with data corrected by all three new corrections, as well as the current correction, were within 0.015 of zero. At the same signal rate in the Omega Centauri data, the current correction produces coefficients closer to -0.09, while the three new corrections result in coefficients closer to -0.05. In this external observation, all of the corrections appear to leave more residual non-linearity than in the case of the flat field. In addition, the improvement when moving from the current correction to the new corrections is much greater than it was in the case of the flat field observation. Similar to the results from the flat field, the 430 correction seems to leave behind slightly lower residual non-linearity than the other new corrections.

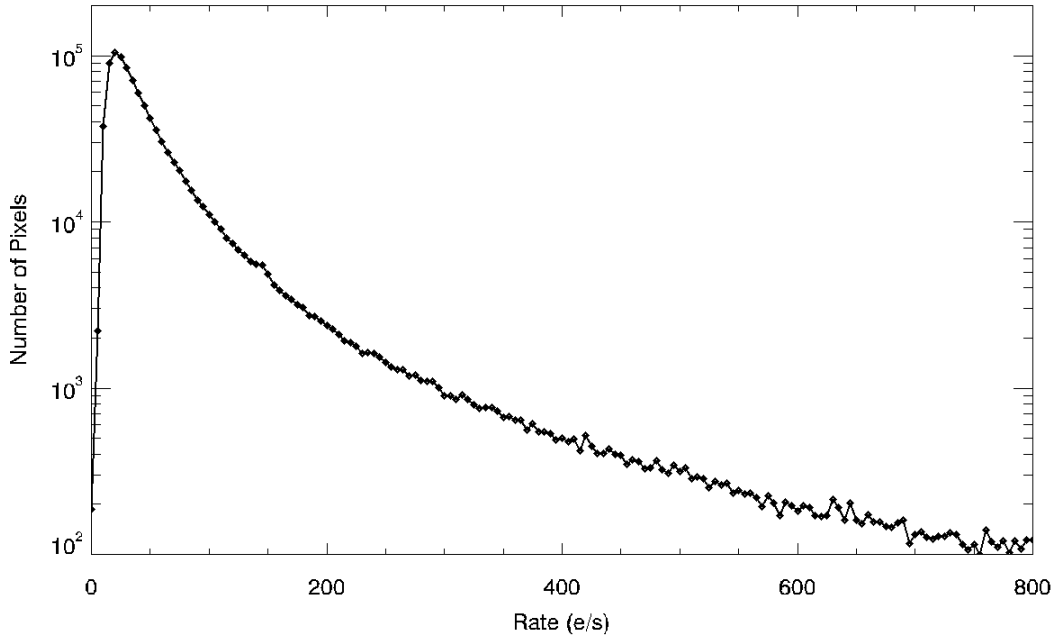


Figure 8: Histogram of initial signal rates for the F160W Omega Centauri ramp. This shows that when placed into rate bins, an acceptable number of pixels were present in bins up to about 800 e-/s.

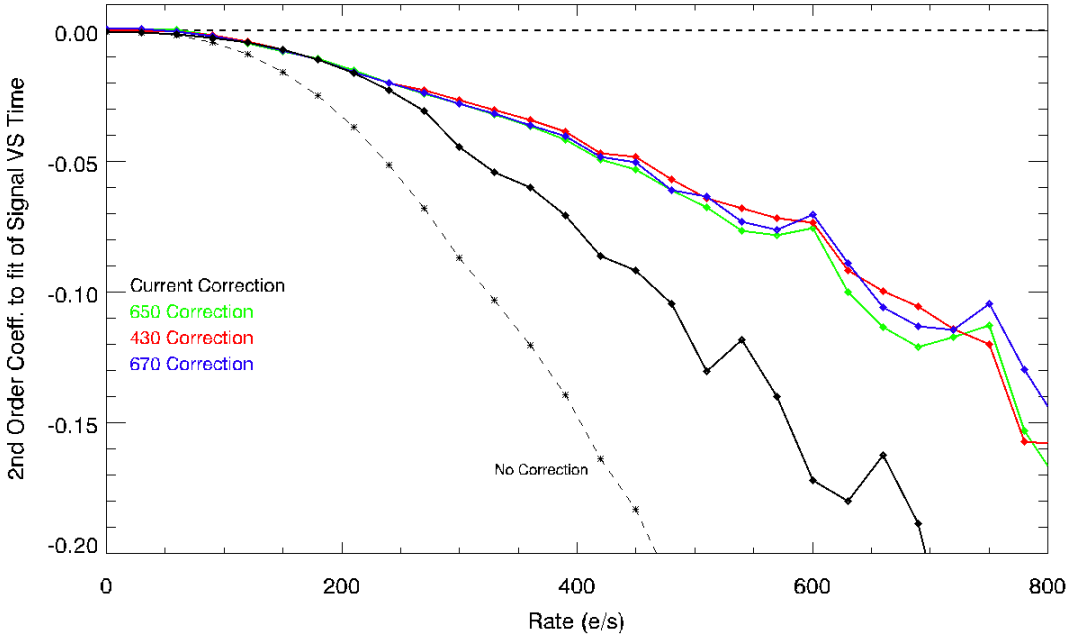


Figure 9: Quadratic coefficients of fits to plots of signal versus time in the corrected ramps. Note the dissimilarity when compared to the same plot for the flat field observation, in figure 6. In this case a higher residual 2nd order contribution remains in the corrected data. There is also a very clear decrease in the magnitude of the residuals for the new corrections compared to the current correction.

In an effort to examine the quality of the various non-linearity corrections using quantities more intuitive than best-fit slopes and quadratic coefficients, we also looked at the percentage differences between the signals up the ramp in the corrected *ima* files and the signals in an "ideal" ramp. The "ideal" signals were calculated by fitting a line to a pixel's measured rate versus time (such as that shown in figure 4). The intercept of this fit was taken as the rate in that pixel at the beginning of the ramp, and declared the "ideal" rate. Multiplying this ideal rate by the exposure time for each read, we arrived at the set of "ideal" signal values for that pixel. We then measured the difference between these ideal signals and those actually present in the ramps after the non-linearity corrections were applied. The percentage difference between the signal in the final unsaturated read and the corresponding "ideal" signal for that read was then recorded.

Note that this method is very similar to the way the non-linearity coefficients were created in the past, with the exception of how the "ideal" ramp was defined. The plot in figure 10 shows the results of these calculations for the Omega Centauri observaiton. There are several features to note in this plot. First, the abrupt change in slope occuring at a rate of 300 e-/sec is an artifact of pixels becoming saturated. All of the files used in the creation of this plot were run through calwf3 which, for each pixel, flagged as saturated any reads where the signal was above that listed in the current saturation map. With an exposure time of 253 seconds, pixels with a signal rate of about 300 e-/sec would have accumulated 76000 e- by the time of their final read. For pixels where this value was flagged as saturated, the second-to-last read was used to calculate the difference plotted below. This second-to-last read, with lower signal than the final read, would also exhibit a lower non-linearity.

Comparing the curves in figure 10, each of the non-linearity corrections lowers the percentage difference between the corrected ramp and it's corresponding "ideal" ramp, compared to the case of no non-linearity correction. At signal rates above 200 e-/sec, all three new corrections give significantly smaller (by up to 2%) residuals than the current correction. Of the three new corrections, the 430 correction again appears to do a slightly better job than the 650 and 670 corrections.

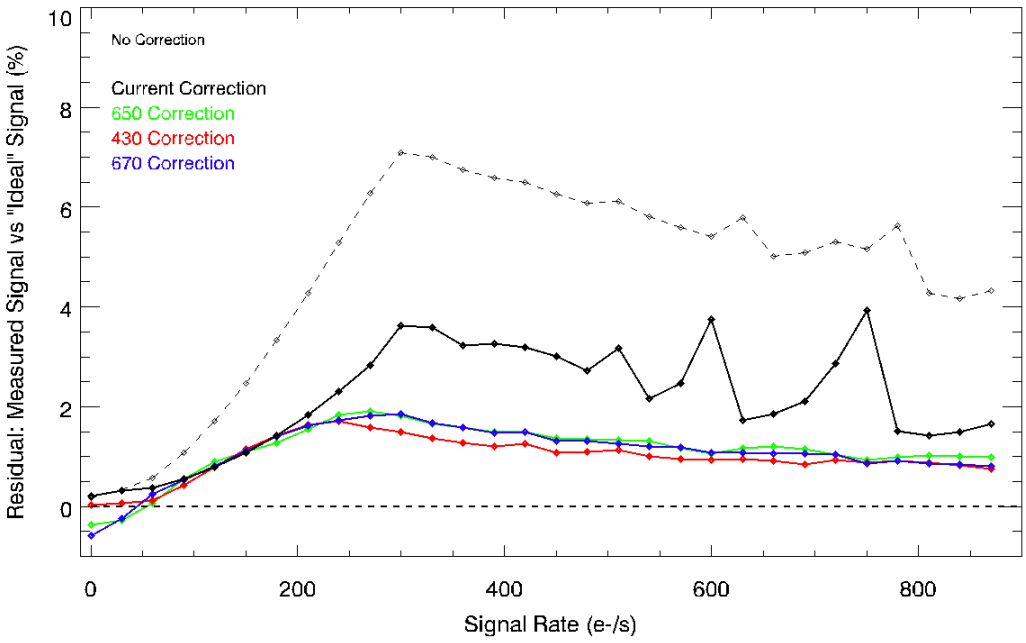


Figure 10: This plot shows the difference, in percent, between the measured signal up the ramp, and the signal in an "ideal" ramp. The new non-linearity coefficients provide smaller residuals than the current non-lineairty coefficients.

Photometry Comparison: Truncated Ramps

We next wished to compare the results of the non-linearity corrections using a method closer to that which would be used on typical WFC3/IR data. We therefore examined the quality and consistency of the various non-linearity corrections through a comparison of photometry on the *flt* files produced from corrected files.

For this test, rather than comparing the photometric results obtained using the various sets of non-linearity coefficients with one another, we focused on comparing how consistent the photometry results were for a given set of non-linearity coefficients. More specifically, did a given set of non-linearity coefficients correct a given source equally well when that source appeared with a bright signal in a long exposure compared to lower signal in a shorter exposure?

In order to assure consistent results, we needed to use observations of the same field of view at the same pointing, but with differing exposure times. This would lead to a given pixel exhibiting different total signal levels from exposure to exposure, but with the same source flux in all exposures. This latter fact meant we could eliminate any rate dependent non-linearity (such as that described by Riess, 2011) as a variable in our analysis. Ideally, after running calwf3 and applying a non-linearity correction, a pixel

seeing the same spot on the sky should have the same value in all of the resulting *flt* files, regardless of exposure time.

Unfortunately, such a dataset does not exist given the dithering and small pointing differences between ramps. In addition, self-induced persistence in back-to-back observations at exactly the same pointing would lead to photometry differences in bright objects for such a dataset. As a result, we decided to create the appropriate data set for this test, using observations of 47Tuc. The files used for this analysis are listed in Table 1 in the appendix.

For each of the 6 files used, the full ramp comprised 16 reads and spanned 352.9 seconds. This complete ramp served as our "long" ramp. We then created a shorter version of the ramp by copying only the first 7 reads over into a new file. This ramp then had an effective exposure time of only 127.9 seconds.

Both the long and truncated ramps were then run through the calwf3 pipeline once using each of the new non-linearity corrections, as well as the current correction, and with no correction. We were then able to compare long and truncated versions for each non-linearity correction.

This comparison was done using very simple photometry on the sources in all of the files. Examination of the results focused on the differences between the short versus long ramps for each non-linearity correction, as well as how these differences differed between non-linearity coefficients. For the photometry we summed the pixel values in the *flt* files using 3 and 5-pixel radius apertures on each source. The results between the two apertures were similar, although the 5-pixel radius aperture produced elevated noise, so we report only on the 3-pixel radius results below.

Figures 11 and 12 show the results from one of the 6 files (ibbw04d6q), using the current non-linearity correction and the 430 correction, respectively. Both plots show the fractional difference of the photometry for the truncated exposure minus the full exposure. Perfect data would appear as a horizontal line at y = 0.

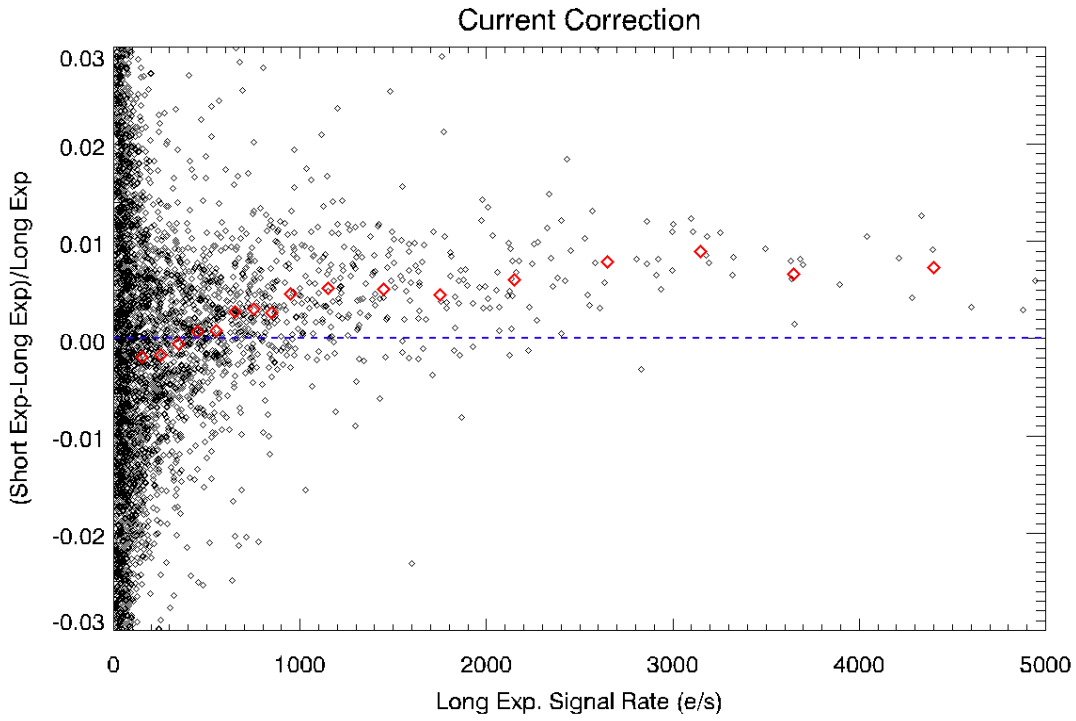


Figure 11: ibbw04d6q, current non-lin correction, long vs truncated ramp photometry. Red diamonds are binned means.

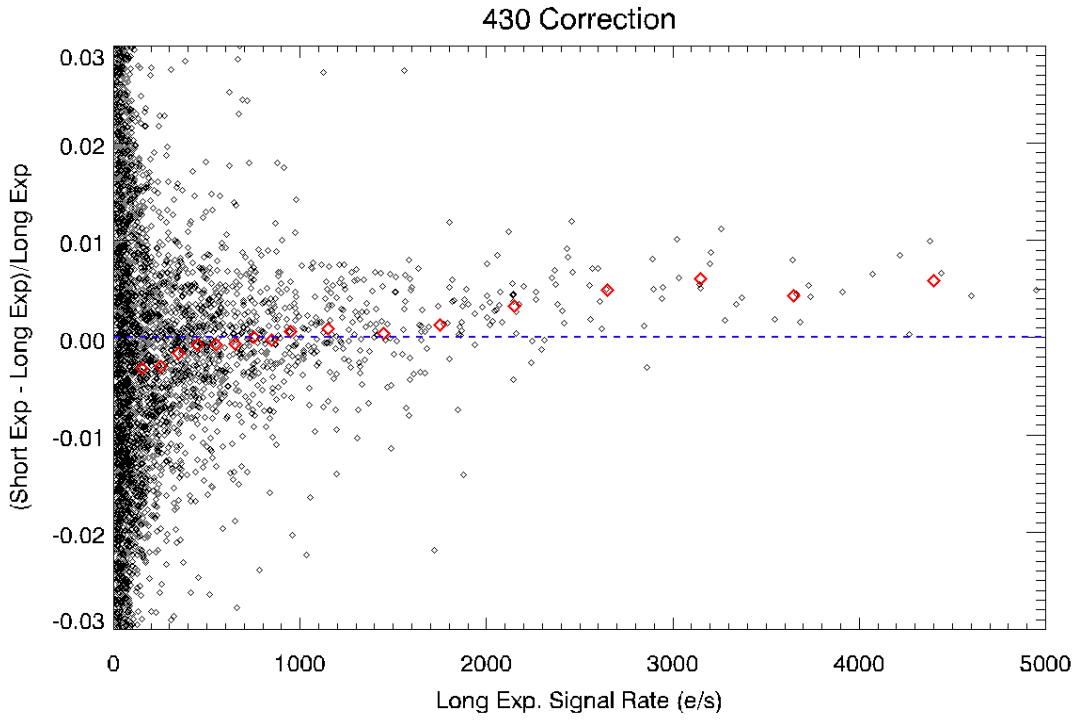


Figure 12: same as above, but using 430 correction. Overall, the red diamonds are have a smaller slope and are closer to zero than when the current non-linearity correciton is used, as seen in figure 11.

In both figure 11 and figure 12, each black point represents the photometry result from a single source in the 47Tuc observation. We used roughly 5600 sources for this analysis. The red diamonds are mean values for binned sets of the black points, plotted to aid the analysis. Fractional differences above zero mean that a given object appears brighter in the short exposure compared to a long exposure. (Keep in mind that this analysis is done on signal rates, not signals. Ideally a source should appear equally bright in the full and truncated exposures after *calwf3* converts them to units of e^-/sec.) A source which is brighter in the truncated exposure than in the full exposure implies that the non-linearity correction undercorrected the signal in the longer exposure, where the non-linearity is a larger effect.

In both figures above we see that the red points slowly increase with increasing measured signal rates, implying that the non-linearity correction leaves behind more residual non-linearity for brighter objects than for dimmer ones. Comapring figure 11 with figure 12, we see that the slope of the red points under the current non-linearity correction is greater than that for the 430 correction, especially for sources with signal rates under about 1000 e^-/sec. In addition, the 430 correction produces points which are closer to zero by up to about 0.005 (or 0.5%) for sources greater than 600 e^-/sec. Both of these results imply that the 430 correction provides a better result than the current non-linearity coefficients. Figures 18 through 20, in the appendix, show similar plots for the 650 and 670 coefficients, as well as the case of no non-linearity correction. We find that the 650 correction produces a plot very similar to that from the 430 correction, but with most points shifted downward by a fraction of a percent. The 670 correction shifts the points downwards by another fraction of a percent. In both cases, objects at the low signal rate end of the plots (rates less than ~600 e^-/sec) are pushed farther below zero. Values below zero imply an object is brighter in the long exposure compared to the truncated exposure. In order for this to occur, the corresponding non-linearity correction must be either overcorrecting these sources in the long exposure, or undercorrecting them in the short exposure. For this reason, we prefer the 430 correction for this file. While not reproduced in this report, plots produced using the other 5 files listed in Table 2 give similar results. The 430 correction consistently provides fractional differences closest to zero when compared to the other sets of correction coefficients.

Photometry Comparison: Mean Photometric Measurements

For our final analysis task, we expanded the breadth of our photometry analysis by using more observations. This analysis was similar to that described above, where we compared photometry results for full ramps versus artificially truncated ramps. In this case, we used the files listed in Table 3 in the appendix. This included 36 full ramps with exposure times of 352.9 seconds, as well as 55 ramps with an exposure time of 93.9 seconds.

As was done in the previous analysis, each file was run through the *calwf3* data reduction pipeline once using each of the new non-linearity coefficients, as well as the current coefficients, and also with no non-linearity correction.

Unlike the observations described in the Truncated Ramps section above, the observations listed in Table 3 all had very similar, but not exactly identical, pointings. As subsequent steps in this analysis involved combining data from all of the observations, we next used astrodrizzle (Gonzaga et al, 2012) to shift all of the input files to be on the same pixel grid. From this point onward, we used the *single_sci* version of the input files for our analysis.

In preparation for photometry, we used source extractor (Bertin & Arnouts, 1996) on one of the long input files to generate an object catalog. With all of the input files now on the same pixel grid, we were then able to use the object positions from this catalog to perform photometry on each of the input files individually.

In an attempt to reduce the noise of the photometry measurements, we then calculated, for each object, the mean signal rate from the group of truncated input ramps, as well as the mean rate from the group of full input ramps. This calculation was done separately for each set of non-linearity correction coefficients.

The figures below show the results of this photometry for the full ramp versus the 93-second ramp. In all figures, we plot the mean photometric results for all of the objects in our catalog.

Figure 13 below shows the fractional difference in photometry when using the current non-linearity correction. In this case, we plot the fractional difference between the photometric results from the 93-second version of the ramp and those from the full 352 second version of the ramp. The x-axis is the photometry results from the full 352 second ramp. This plot shows that the vast majority of the sources in the field were consistent to better than 1%. The bulk of the lowest-flux objects appear to be ~0.5% brighter in the short ramp compared to the full ramp. The overall shape of this population of points is fairly similar to that seen in figure 11. The main difference appears at the lowest source fluxes. While the population of low-flux objects appears centered a few tenths of a percent below zero in figure 11, the population appears several tenths of a percent above zero in figure 13 below.

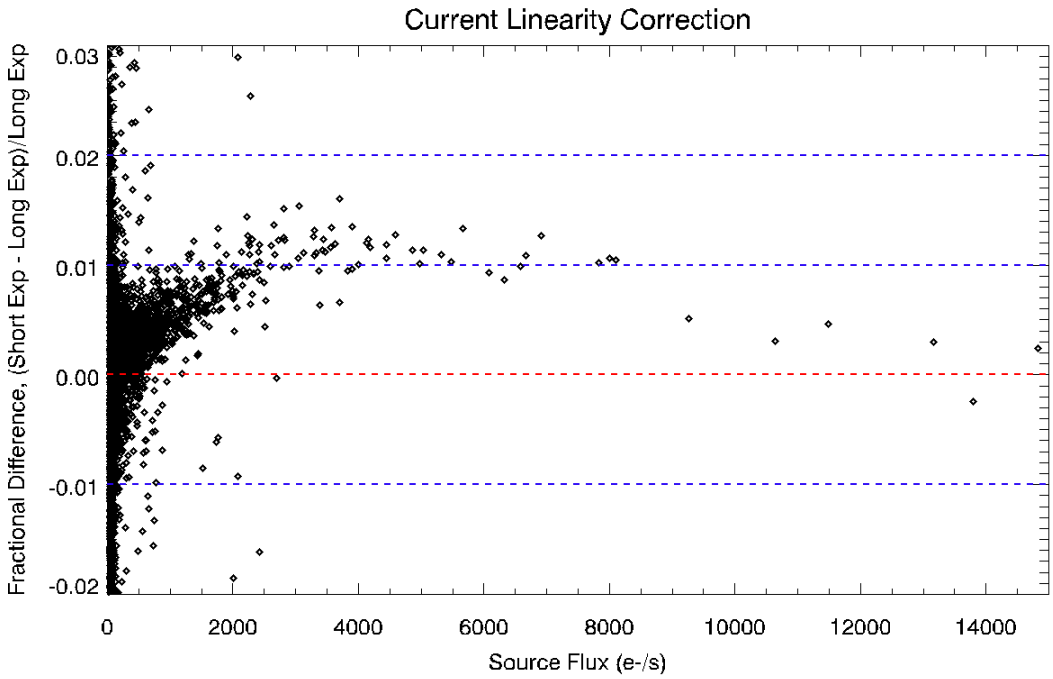


Figure 13: Photometry results for the mean of each object in our catalog. This plot shows the difference of the measured signal rate in the 93-second, truncated ramps to that in the 352 second, full ramps. This difference is normalized by the signal rates in the full ramps.

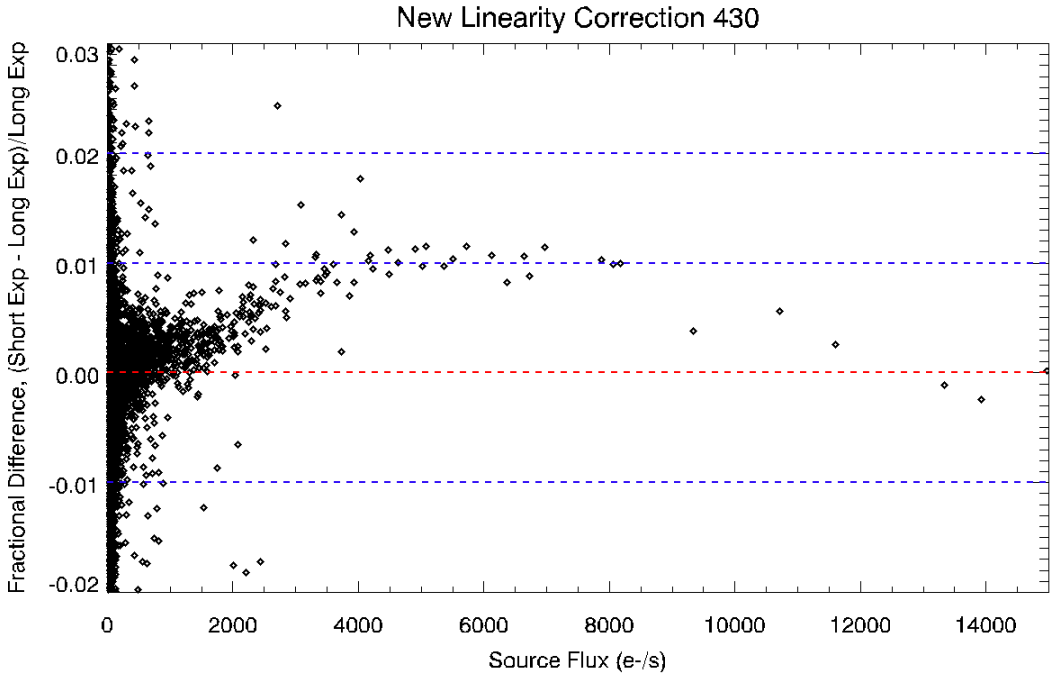


Figure 14: Same plot as figure 13 above, but for files corrected with the new 430 coefficients.

Figure 14 is the same plot seen in figure 13, but for files corrected with the new 430 correction coefficients. The overall behavior seen in this plot is similar to that in figure 13. The low-flux objects (those with flux < 2500 e-/sec) appear slightly closer to zero than when the current non-linearity correction is used. Many of the points with signal rates under 3000 e-/s and large fractional differences (large *y* values) in figure 11 have been pushed down much closer to zero in the plot above. The large-scale, sinusoidal-like shape of the points is still present in this version of the plot. Overall, the shape of the population is very similar to that in our initial photometry comparison, seen in figure 12.

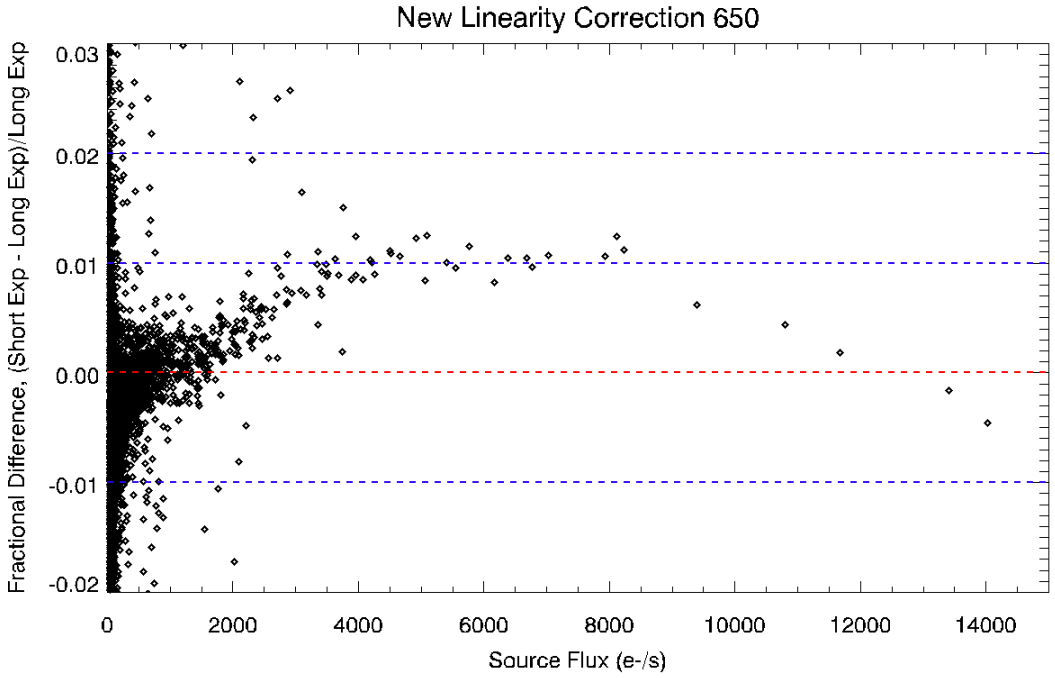


Figure 15: Same as figures 13 and 14 above, but for files corrected with the new 650 coefficients.

The photometry results from files created with the new 650 linearity coefficients are very similar to those from the 430 correction. Figure 15 shows that when the 650 coefficients were used, the population of points with fluxes below 2500 e-/s has been pushed slightly closer to zero compared to that in figure 14. At higher fluxes, the photometry results are nearly identical to those from the 430 case.

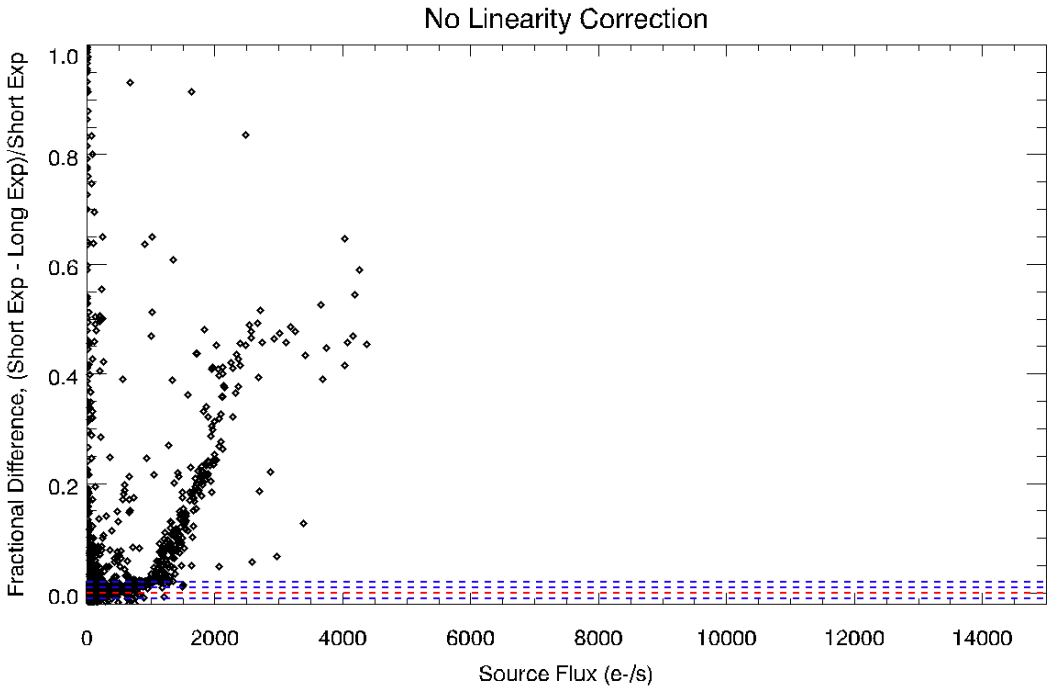


Figure 16: This plot shows the comparison case, where no linearity correction was made. The x-scale has been kept the same for ease of comparison with figures 13 through 15.

Finally, for completeness, we show in figure 16 a similar plot, for the case of no linearity correction. We have kept the same horizontal range, as well as dashed blue and red lines as in the previous plots, in order to clearly show the large differences in photometry that result from uncorrected data.

Conclusions

Based on the the analysis presented above, we believe that this updated method of non-linearity correction calclation is superior to that used previously for the WFC3/IR channel. Measurements of residual non-linearity as well as photometry on corrected data show that the new coefficients are able to better remove non-linearity than the correction currently in use.

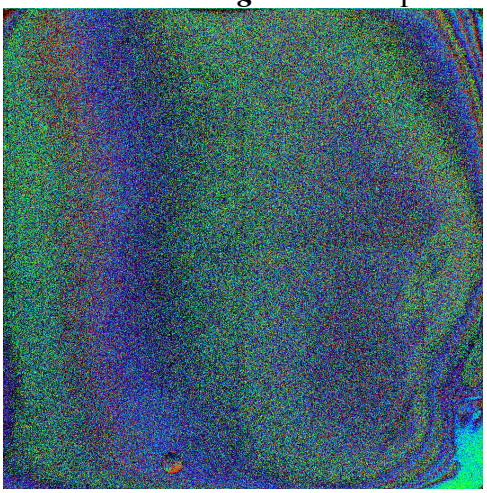
Among the three sets of new correction coefficients compared here, we believe that the 430 coefficients provide the best overall correction, although the difference between them and the 650 coefficients are small.

Figure 17 below gives a graphical representation of the α, β, γ, and δ coefficients from equation 5 for the 430 case. These maps show significant spatial variation in the coefficients across the detector, which match well the structure seen in on-orbit internal flat field images (see figure 6 in Dahlen, 2013). This is not completely unexpected, as the

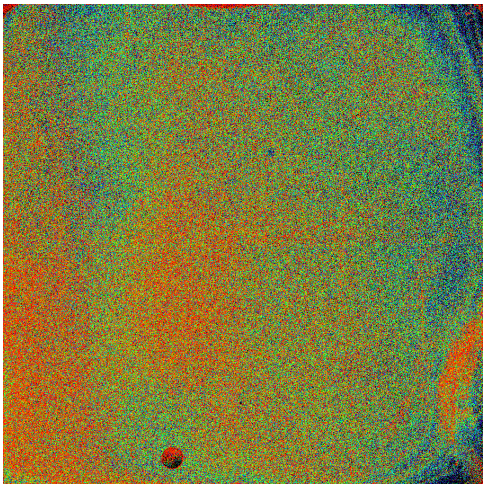
correction coefficients were derived from flat field images, and the non-linearity correction is performed prior to the flat fielding step during the data reduction process.

In the current non-linearity correction, the *calwf3* pipeline uses only a single set of coefficients for each quadrant of the detector. The spatial variations seen below suggest that using a separate set of coefficients for each pixel provides a better non-linearity correction. We did not investigate the middle ground between these two options. Future work may indicate whether averaging or binning smaller numbers of pixels can improve the correction further.

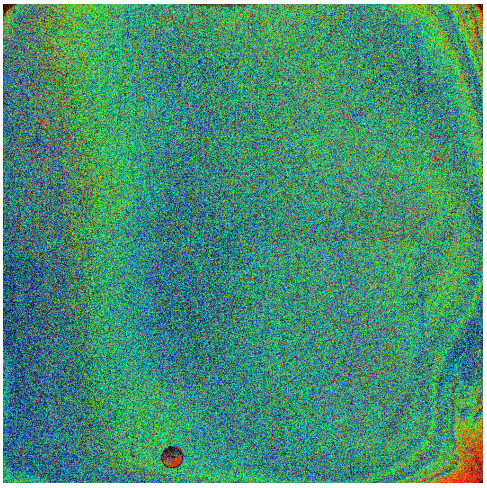
Figure 17: Maps of the 430 correction coefficients.



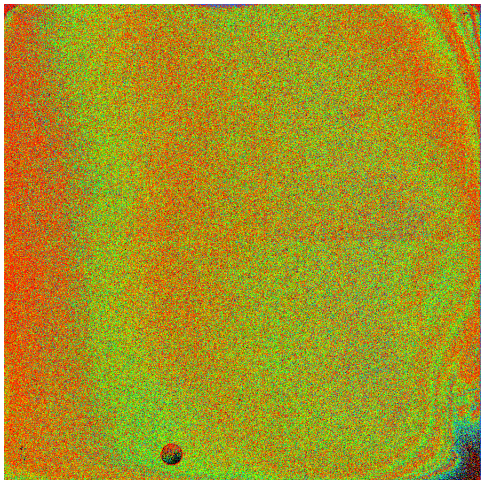
0^{th} order coeff (α). Linear stretch, from -0.002 to +0.002.



Slope coefficient (β), linear stretch -0.5×10^{-6} to 1.5×10^{-6}.



Quadratic coefficient (γ), linear stretch -1×10^{-10} to 0.5×10^{-10}.



3^{rd} order coefficient (δ), linear stretch -1×10^{-15} to 2×10^{-15}.

We also note that for all correction coefficients, there is a significant difference in the quality of the non-linearity correction for flat fields versus point sources. Flat field data

appears to have much more of the non-linearity removed compared to the external, point-source observations. This may argue for a non-linearity correction derived from point-source data in the future.

Recommendations

Based on the results presented here, we recommend updating the non-linearity correction for WFC3/IR to the 430 coefficients. Note however that in order to keep the flux calibrations correct, the IR channel zeropoints would need to be updated at the same time.

References

Bertin, E. & Arnouts, S. 1996: SExtractor: Software for source extraction, Astronomy & Astrophysics Supplement 317, 393.

Dahlen, T., WFC3 ISR 2013-04. "WFC3/IR Internal Flat Fields"

Gonzaga, S., Hack, W., Fruchter, A., Mack, J., eds. 2012, The DrizzlePac Handbook. (Baltimore, STScI)

Hilbert, B., WFC3 ISR 2008-39. "WFC3 TV3 Testing: IR Channel Nonlinearity Correction"

Markwardt, C. B. 2009, "Non-Linear Least Squares Fitting in IDL with MPFIT," in proc. *Astronomical Data Analysis Software and Systems XVIII*, Quebec, Canada, ASP Conference Series, Vol. 411, eds. D. Bohlender, P. Dowler & D. Durand (Astronomical Society of the Pacific: San Francisco), p. 251-254 (ISBN: 978-1-58381-702-5;)

Rajan, A. et al. 2010, "WFC3 Data Handbook", Version 2.0, (Baltimore:STScI).

Riess, A. WFC3 ISR 2011-15. "An Independent Determination of WFC3-IR Zeropoints and Count Rate Non-Linearity from 2MASS Asterisms"

Robberto, M., JWST-STScI-CR-002163, July 2010, "An improved algorithm for the correction of IR detector non linearity".

Robberto, M., JWST-STScI-CR-002346, January 2011. "Implementation of a new algorithm for the correction of non-linearity in JWST detectors".

Robberto M. and Hilbert, B. WFC3 ISR 2005-29. "WFC3 2004 Thermal Vacuum Campaign: IR channel linearity (flat field illumination - SMS IR04)"

Acknowledgements

Many thanks to Massimo Robberto for supplying the core of this updated correction method, as well as numerous discussions of the new technique, and also to Gabe Brammer for a thorough and helpful review of this paper.

Appendix

Files used to create the mean flat field ramp that was used to calculate the new non-linearity coefficients are listed in Table 1 below.

Ramp Name	Proposal ID	Filter	Sample Sequence	Exposure Time (sec)
ibmg01itq, ibmg02sfq, ibmg03a7q, ibmg04ceq, ibmg05dcq, ibmg06eaq, ibmg07jkq, ibmg10uvq, ibmg13gsq, ibmg14q3q, ibmg15qnq, ibmg16r9q	12352	Blank	SPARS25	352.9
ibmg01iwq, ibmg02siq, ibmg03aaq, ibmg04chq, ibmg05dfq, ibmg06edq, ibmg07jtq, ibmg10uzq, ibmg13gvq, ibmg14q6q, ibmg15qqq, ibmg16rcq	12352	F127M	SPARS25	352.9
ibvl01adq, ibvl02dqq, ibvl03f2q, ibvl04kbq, ibvl05o8q, ibvl13soq, ibvl14ovq, ibvl15fxq, ibvl16k2q	12696	Blank	SPARS25	352.9
ibvl01ahq, ibvl02dtq, ibvl03f5q, ibvl04khq, ibvl05obq, ibvl13suq, ibvl14p2q, ibvl15g0q, ibvl16k6q	12696	F127M	SPARS25	352.9
ibvl06slq, ibvl07voq, ibvl18v5s, ibvl19v2q	12696	Blank	SPARS25	352.9
ibvl06snq, ibvl07vtq, ibvl18v7s, ibvl19v4q	12696	F127M	SPARS50	452.9
ic5n07jlq, ic5n08khq, ic5n09mwq	13079	Blank	SPARS25	352.9
ic5n07joq, ic5n08kkq, ic5n09mzq	13079	F127M	SPARS25	352.9

Table 1: *Files used to create the non-linearity correction coefficients.*

File	Proposal ID	Filter	Sample Sequence	Exposure Time (sec)
ibbw04d6q	11931	F160W	SPARS25	352.9
ibbw04d8q	11931	F160W	SPARS25	352.9
ibbw08muq	11931	F160W	SPARS25	352.9
ibbw08mwq	11931	F160W	SPARS25	352.9
ibmg20k9q	12352	F160W	SPARS25	352.9
ibmg21gfq	12352	F160W	SPARS25	352.9

Table 2*:* Observations used for the comparison of full versus truncated ramp photometry. All observations were of 47Tuc. In order to create the truncated ramps, we removed the final 9 reads, resulting in ramps with effective exposure times of 127.9 seconds.

File	Proposal ID	Sample Sequence	Exposure Time (sec)
ibbw04d5q,ibbw04daq,ibbw04dcq, ibbw04dhq,ibbw04djq,ibbw04doq ibbw04dqq,ibbw04dvq,ibbw08mtq ibbw08myq,ibbw08n0q,ibbw08n5q ibbw08n7q,ibbw08ncq,ibbw08neq ibbw08njq,ibbw09msq,ibbw09mxq ibbw09mzq,ibbw09n4q,ibbw09n6q ibbw09nbq,ibbw09ndq,ibbw09niq	11931	SPARS10/11 reads	92.9
ibmg20k6q,ibmg20kdq,ibmg21gcq ibmg21glq,ibmg22pzq,ibmg22q6q ibmg22q7q,ibmg23c4q,ibmg23crq ibmg24wbq,ibmg24wiq,ibmg25kuq ibmg25l1q,ibmg25l2q	12352	SPARS10/11 reads	92.9
ibvl10h0q,ibvl10hrq,ibvl11f7q ibvl11fyq,ibvl12kdq,ibvl12kkq ibvl12klq,ibvl22lbq,ibvl22liq ibvl23p0q,ibvl23p7q,ibvl24xvq ibvl24yaq,ibvl24ybq	12696	SPARS10/11 reads	92.9
ic5n1es1q,ic5n1es8q,ic5n2em9q ic5n2en2q	13079	SPARS10/11 reads	92.9
ibbw04d6q,ibbw04d8q,ibbw04ddq ibbw04dfq,ibbw04dkq,ibbw04dmq ibbw04drq,ibbw04dtq,ibbw08muq ibbw08mwq,ibbw08n1q,ibbw08n3q ibbw08n8q,ibbw08naq,ibbw08nfq ibbw08nhq,ibbw09mtq,ibbw09mvq ibbw09n0q,ibbw09n2q,ibbw09n7q ibbw09n9q,ibbw09neq,ibbw09ngq	11931	SPARS25/16 reads	352.9
ibmg20k9q,ibmg21gfq,ibmg22q2q ibmg23caq,ibmg24weq,ibmg25kxq	12352	SPARS25/16 reads	352.9
ibvl10hnq,ibvl11fuq,ibvl12kgq ibvl22leq,ibvl23p3q,ibvl24y2q	12696	SPARS25/16 reads	352.9
ic5n1es4q,ic5n2emcq	13079	SPARS25/16 reads	352.9

Table 3: 47Tuc observations used in the final photometry comparison.

Photometry results for ibbw04d6q:

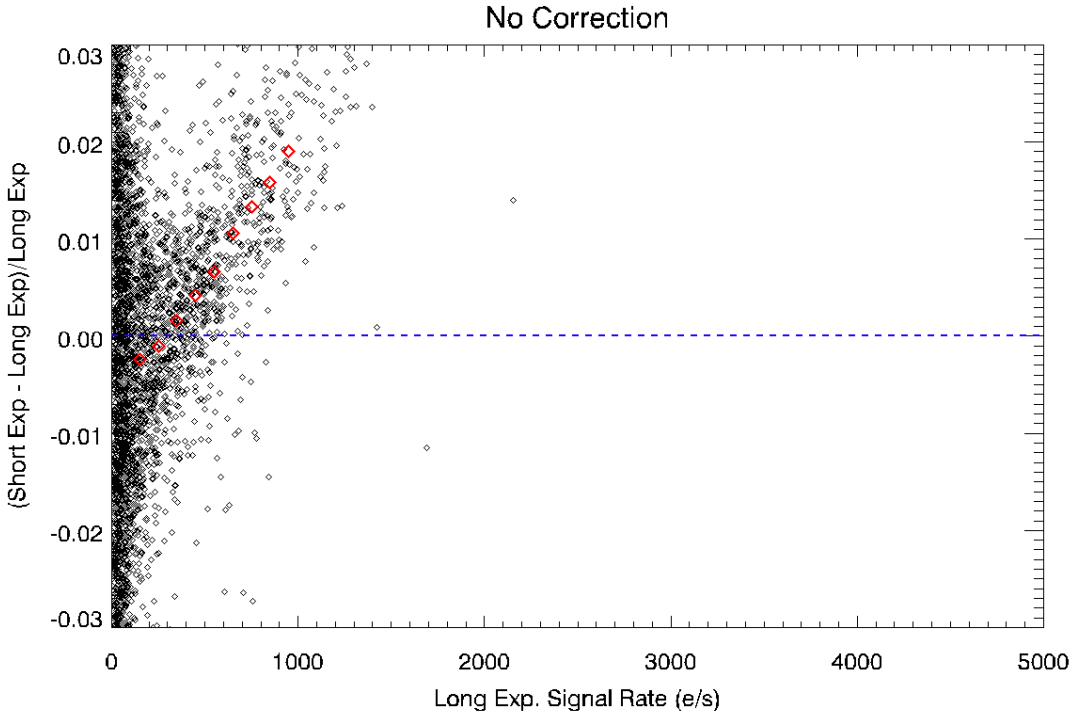


Figure 18: Same plot as seen in figure 12, but for the case of no non-linearity correction applied. In this case, photometry results in a given object in the truncated ramp are measured to be brighter than in the full ramp. This is because the decreasing signal rate with time (as seen in the top panel of Figure 1), leads to a shallower best-fit slope in the full ramp where more signal accumulated in the longer exposure time.

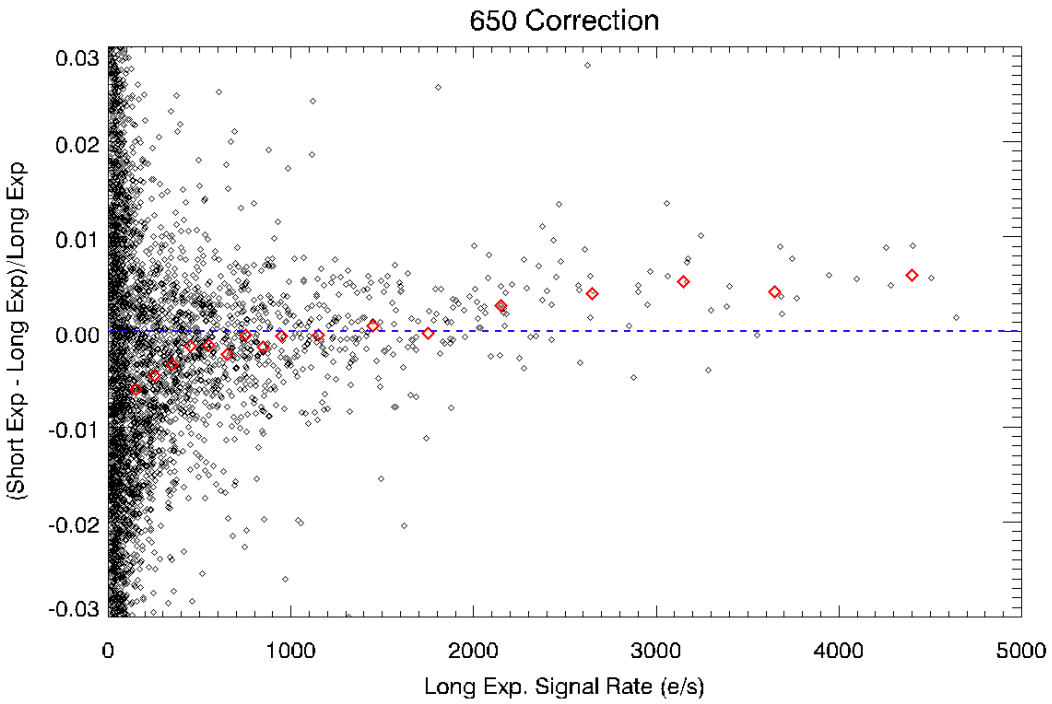


Figure 19: The same plot as figure 12, but with the 650 correction applied to the full and truncated ramps. Overall, the population of black and red points has been pushed downward by a fraction of a percent compared to that from the 430 correction.

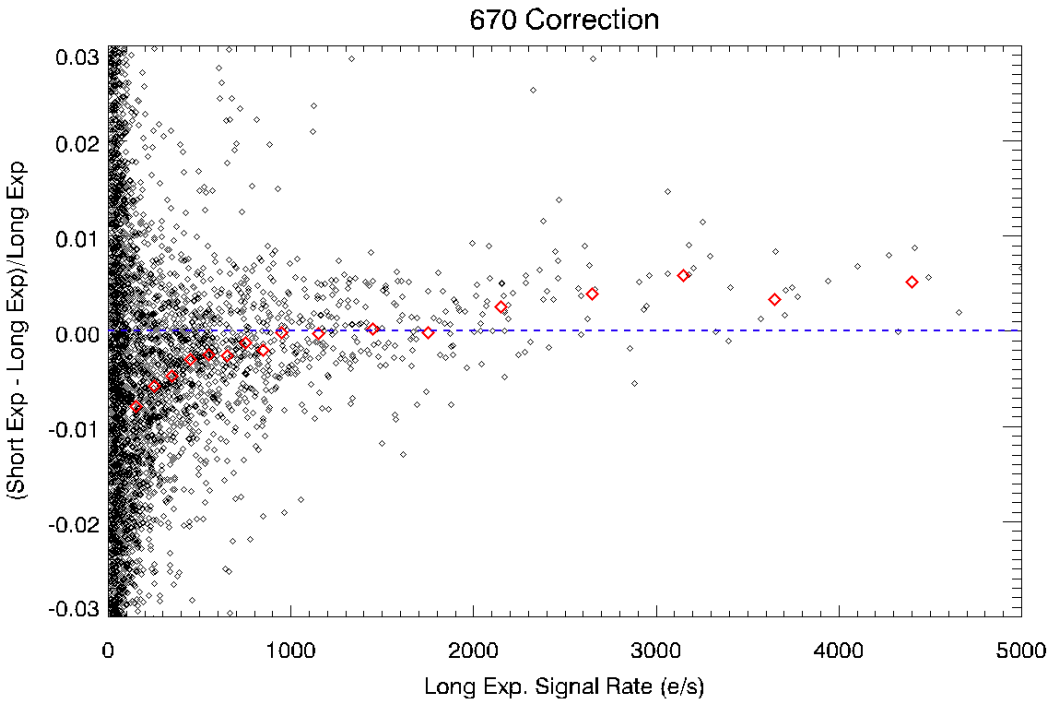


Figure 20: The same plot as figure 12, but with the 670 correction applied. This correction pushes the population of points down even further than the 650 correction above. Points below zero show a case where an object is brighter in the long exposure than in the truncated exposure, implying an overcorrection of the non-linearity in the long exposure.