WHIRC Report VI – Mean-Variance Analysis

Dick Joyce
5 January 2010; revised 28 October 2010

1. Fowler-Mortara Mean-Variance Analysis

Modern electronic detectors are read out in a digital/analog/digital process, in which the electrons are collected on a capacitor (usually a potential well produced by reverse bias of a diode junction), read out as an analog voltage, then converted again to a digital output by an analog-to-digital converter. The relation between the physical signal in electrons and the "arbitrary digital units" (ADUs) which comprise the observed data can be determined by an analysis of the noise as a function of the signal level. This analysis, generally referred to as a Photon Transfer Curve (PTC) or mean-variance analysis, involves generating a dataset of images taken at signal levels ranging from near zero (darks) to the saturation level of the detector, and calculating both the mean level and standard deviation of the mean of an ensemble of images at each signal level.

Under the assumption that the only sources of noise are the zero-signal (read) noise and the Poisson noise of the photon (or dark current) signal, Mortara and Fowler (1981, *SPIE, 290*, 28) defined the variance of the image (in electrons) as

$$V_e = N_{re}{}^2 + S_e \quad [1]$$

Where N_{re} and S_e are the read noise and signal, in electrons.

If g is the conversion gain in e/ADU, then

$$g^2 V_{adu} = g^2 N_{radu}{}^2 + g S_{adu} \quad [2]$$

$$S_{adu} = g\,(V_{adu} - N_{radu}{}^2) \quad [3]$$

$$\log\,(V_{adu} - N_{radu}{}^2) = \log\,(S_{adu}) - \log\,(g) \quad [4]$$

By measuring the standard deviation and mean signal in ADU and determining the read noise from the zero-signal noise, one can plot the read-noise subtracted variance against the mean signal to obtain a linear curve of slope=1.0, whose X intercept is the gain.

If one writes equation [3] as

$$V_{adu} = N_{radu}{}^2 + g^{-1}\,S_{adu} \quad [5]$$

Then one should get a linear plot whose Y intercept is $N_{radu}{}^2$ and slope is 1/g.

Although both techniques should give the same result, the author prefers the logarithmic analysis because it permits one to display a greater range of data and the requirement that the fit *must* have a slope of 1.0 provides a discriminant against flaky or unreliable data. Because the large signal/variance data points exert significant leverage on the Y-intercept when using the linear analysis, one must "anchor" the fit using the measured read noise variance from dark frames as the Y-intercept.

2. Models

Unfortunately, real detectors complicate this analysis because they contain various artifacts such as pixels which are dead, unusually noisy, or variable in sensitivity. It can be difficult to differentiate pixel sensitivity variations due to quantum efficiency from those due to node capacitance. The first of these usually results from variations in the efficiency of the anti-reflection coating on the detector, whereas the second manifests in a real difference in the conversion gain.

To test the techniques of mean-variance analysis, we generated images using the IRAF task 'mknoise', which allows one to specify the conversion gain, read noise, signal level in ADU, and whether to include the Poisson noise from the signal into the image. In keeping with the nominal WHIRC characteristics, the gain was set to 4.0 e/ADU and the read noise to 40 e. Signal levels from 0 to 50000 ADU were used to generate images, and ten images at each signal level were randomly generated to simulate the technique used for obtaining PTC data with WHIRC.

2.1 Uniform Images

In the first case, the images produced by 'mknoise' were uniform, having the same mean level over the entire 2048×2048 image. The mean and standard deviation were calculated in the same way as with WHIRC data, using the IRAF task 'imcomb' to calculate mean and sigma images (in the latter, the value of each pixel is the standard deviation of the mean of the values for that pixel in the input images). Since the only noise, by definition, was the read and Poisson noise, no rejection was employed. The results are shown in Figure 1, where the data are plotted both logarithmically and linearly.

In the logarithmic plot, the subtraction of the read noise variance yields a linear plot of unity slope, but the X intercept is closer to 4.2 than the 4.0 which was the input to the IRAF task. The Y intercept is at 94.6. The linear plot to the total variance yields a good fit to a linear plot with a slope of 0.236 and a Y intercept of 94.7. This translates to g = 4.23 and a read noise of 41.2 e. The two results are, as expected, the same. However, there is clearly something about the way in which the 'imcomb' task calculates the sigma images which results in a gain which is different from the expected value.

As an experiment, we looked at the mean images from the 'imcomb' task and calculated the *spatial* statistics using 'imstat'. Again applying both the logarithmic and linear approaches (Figure 2), we now get values for the read noise variance (99.9 and 100.3,
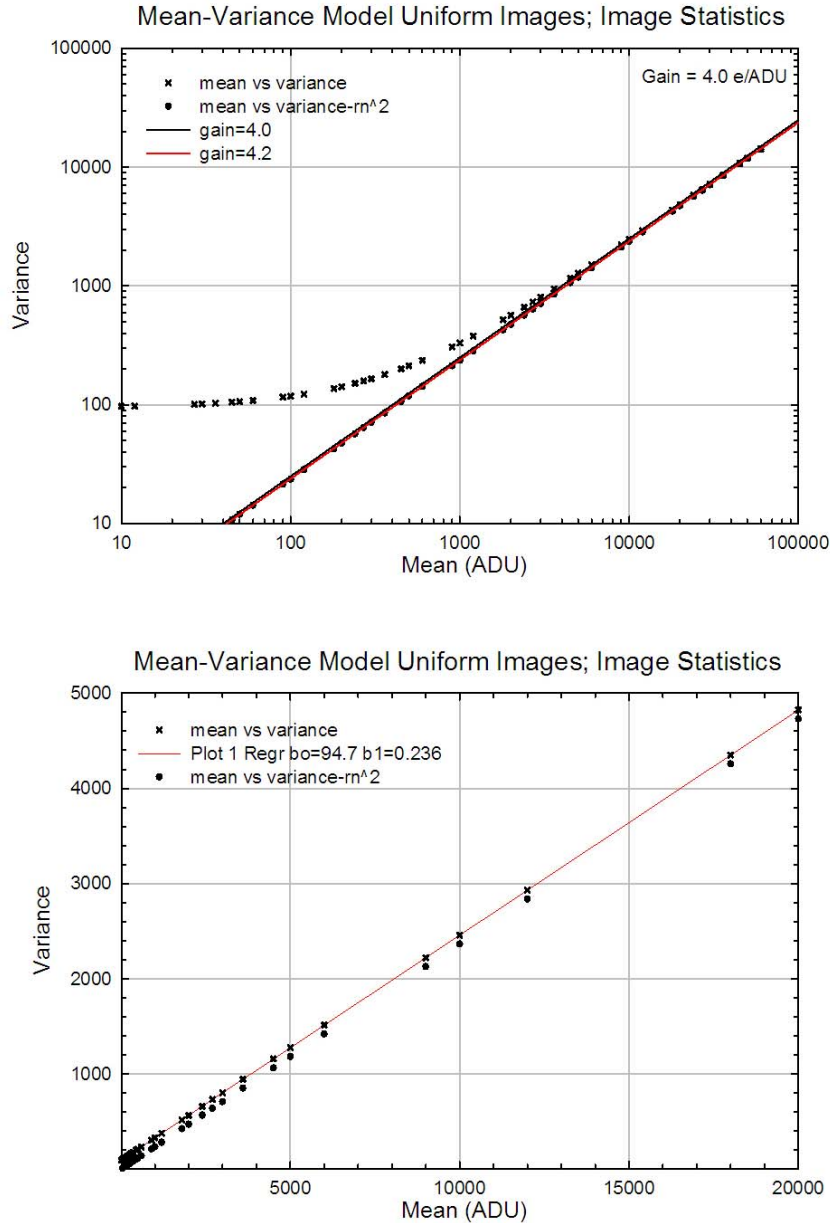
Figure 1: **Logarithmic (top) and linear (bottom) plots of the mean-variance analysis for model images generated using the IRAF task 'mknoise'. The variance was determined on a pixel-by-pixel basis, using the IRAF task 'imcomb' to generate mean and sigma images from 10 independently generated images. The gain was defined to be 4.0 e/ADU, the read noise 40 e, with constant background levels ranging from 0 to 50000 ADU. The only sources of noise are the read noise and the Poisson noise corresponding to the uniform background. In all of these plots, the x symbols represent the variance of the total noise and the dots represent the variance with the square of the read noise subtracted. Note that the gain derived from this analysis is slightly larger than the input parameter to the IRAF task.**
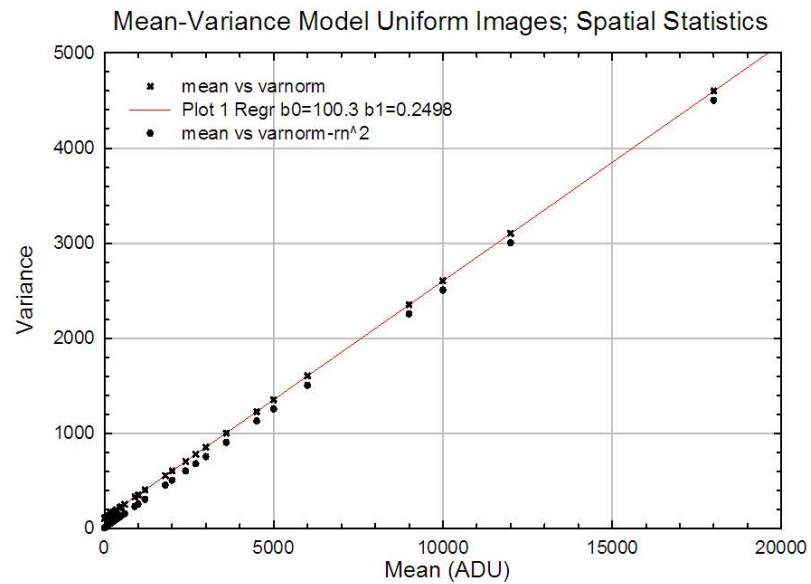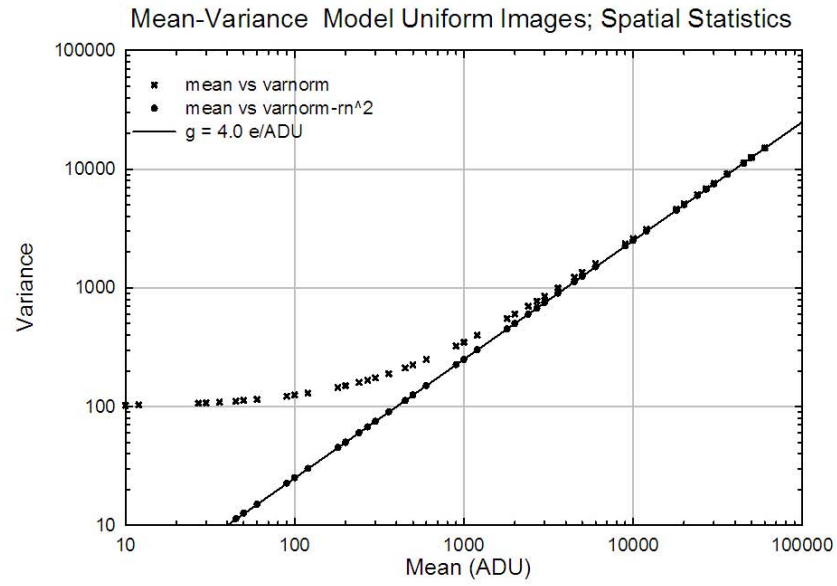
Figure 2: Same as Figure 1, except that the variance was generated from the spatial statistics of the mean image. Note that this gives a fit which is closer to the true gain value of 4.0, but implicitly assumes that the image is precisely flat.

respectively) and gain (4.0 for both) which are the anticipated values. The same number of pixels is involved in both approaches; in the first, one generates a sigma image from 10 individual images and averages over the 4×10^6 pixels of that image; in the second, one averages the 10 images and calculates the sigma over the 4×10^6 pixels. From the viewpoint of the astronomer, the latter is more consistent with determining the limiting performance, since one is attempting to detect a source against a spatial structure of noise. However, the approach assumes that the underlying response of the detector be uniform, which is almost never the case in the real world.

2.2 Nonuniform Images

To simulate the variation in sensitivity which is seen with real detectors, we generated "mixed" images consisting of a combination of different signal levels. For example, the uniform 10000 ADU images consisted of ten independently derived 'mknoise' images with a mean level of 10000 ADU. The "mixed" 10000 ADU images required the generation of ten independent 'mknoise' images with mean levels of 9000, 10000, and 12000 ADU (each with its level of Poisson noise). These were copied into the final images in the respective proportion of 15%, 80%, and 5% to mimic the sensitivity falloff at the top and bottom of the WHIRC array, as well as the elevated pupil ghost in the center. These "mixed" images were then analyzed in the same way as the uniform images.

The approach of calculating statistics on a pixel-by-pixel basis gave results (Figure 3) which were almost identical to those for the uniform images. One can see, particularly in the linear plot, that the mean values are slightly lower (the mean for the 10000 ADU mixed image is 9951 ADU), but the fit to the data in both approaches yields essentially the same result for the gain and read noise. This suggests that small variations from the mean value do not greatly affect the analysis, provided the variations are a result of photon collection efficiency (antireflection coating variations, pupil ghost). If the conversion gain varied significantly over the array, then characterizing the entire array with a single gain would not be appropriate.

Calculating statistics of the mixed images using spatial statistics failed miserably, as one might expect, since the variance was dominated completely by the spatial structure of the mean value rather than the Poisson statistics of the signal (Figure 4).

Mean-Variance Model Mixed Images; Image Statistics

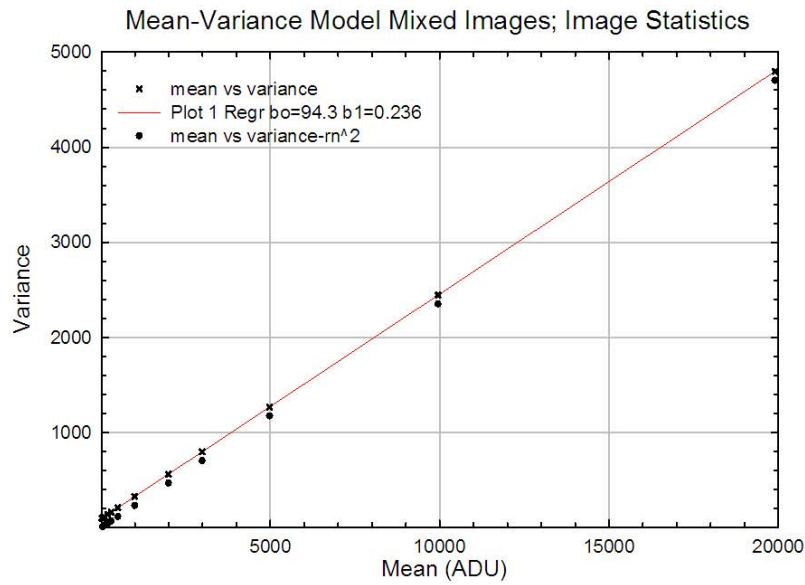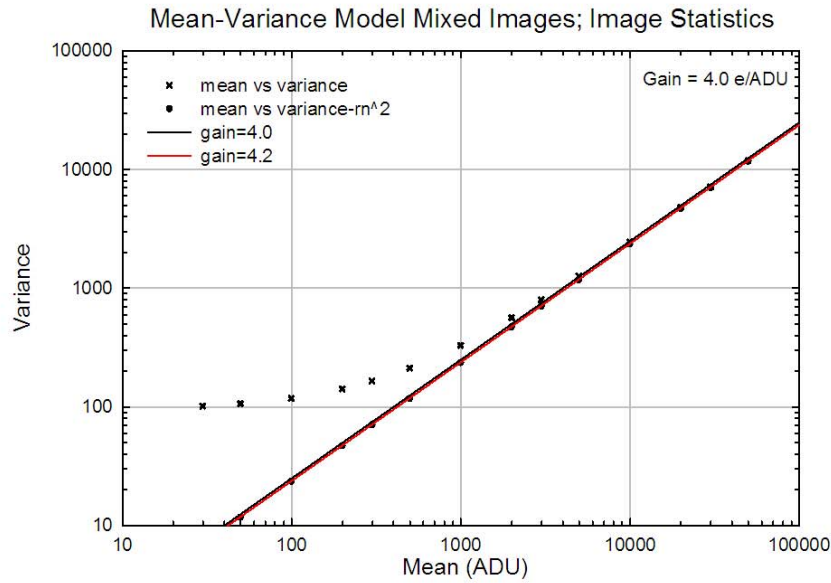Mean-Variance Model Mixed Images; Image Statistics

Figure 3: Same as Figure 1, except that each image consists of a mix of pixels at the nominal ADU level (80%), 10 percent below the nominal level (15%), and 20 percent above the nominal level (5%). This was to simulate the lower response at the top and bottom of the array, as well as the central pupil ghost seen in WHIRC images. As expected, the mean value for each image is slightly less than that in Figure 1, but the gain determined from the analysis is essentially identical.
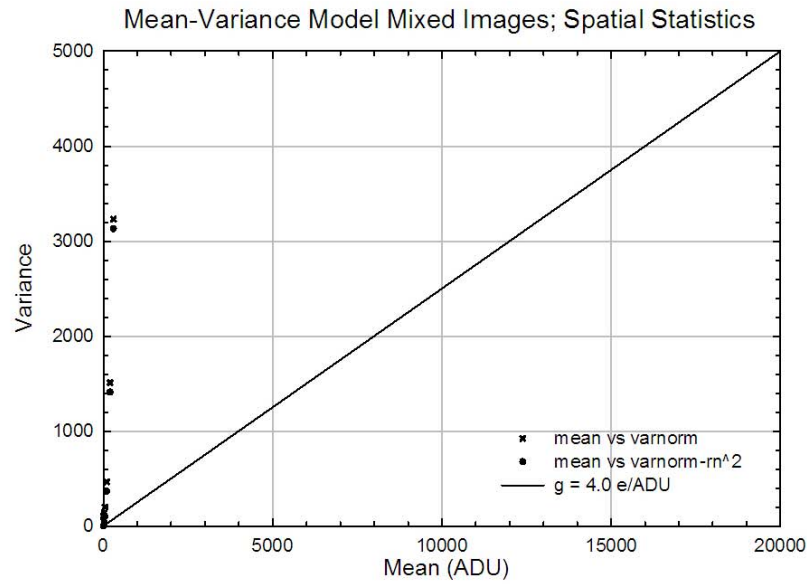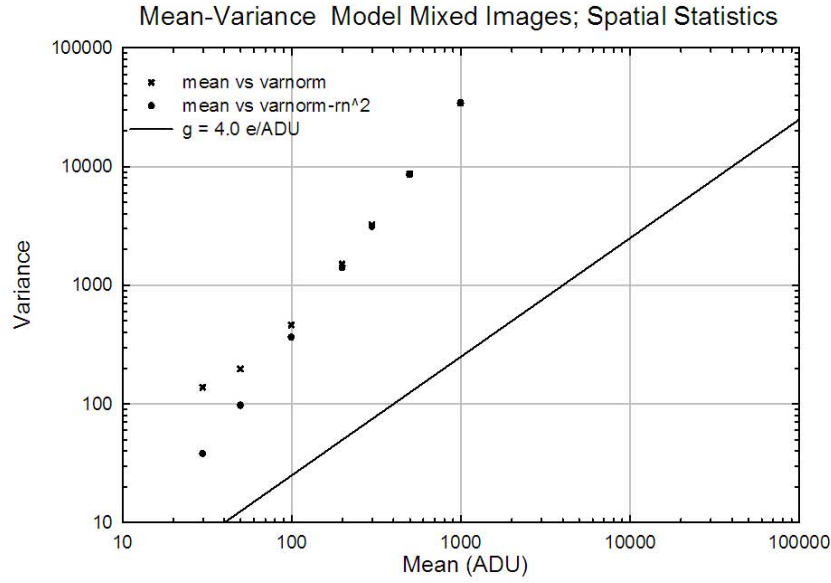
Figure 4: Same as Figure 3, except that the statistics were determined from the spatial statistics of the mean image, as in Figure 2. Note that the mix of mean pixel values completely dominates the variance and makes the mean-variance analysis impossible.

2.2.1 Difference Analysis

Schubnell *et al*. (2006, *SPIE*, **6276**), in their report on the analysis of the infrared detectors for SNAP, noted that the conversion gain can be calculated even for images with spatial variation by taking two images with the same illumination and calculating the difference and sum of the two images. The difference image should have a mean value of 0.0, but the noise will be that of the sum of the two images:

$$V_e = 2N_{re}^2 + S_{e1} + S_{e2} \quad [6]$$

This is identical to equation [1], except that the zero-signal variance is twice the square of the read noise. Since the spatial variations due to illumination variations should be eliminated from the difference image, one may use spatial statistics (which involve a large number of pixels) to calculate the sigma of the difference image and the mean of the sum image, which can then be analyzed using either the linear or logarithmic method.

To evaluate this, we analyzed the "mixed" images described above, calculating the difference and sum of only two images for each illumination level. The results, plotted in Figure 5, demonstrate that this technique yields reliable results.

As an extreme case, we generated "supermixed" images in which half of each image consisted of the next highest illumination level (e.g., half at 10000 ADU, half at 20000 ADU). These dramatically nonuniform images nonetheless yielded the correct values for the gain and read noise (Figure 6).

3. Conclusions from Model Images

- For the model data, the logarithmic and linear approaches give the same results, as one might expect. The author will argue that the logarithmic approach, by requiring the slope to be 1.0, offers a better diagnostic of spurious data in real images, which might simply yield a different slope (and gain) in the linear approach. However, unreliable data in the linear approach would probably also yield an unreasonable Y intercept.
- It is not clear why the image averaging with 'imcomb' appears to give a sigma image which is less than that which is obtained by the spatial averaging, resulting in a calculated gain (by either technique) which is too high.
- The approach suggested by Schubnell *et al*. (which the author admittedly never thought of) appears to be worthy of consideration. The use of spatial statistics involves a much larger number of pixels and should give a better fit with real data. Calculating statistics on the difference image should eliminate sensitivity variations across the array and subtract out dark current.
- All of the mean-variance analysis techniques assume that the only sources of noise are the read noise and signal-dependent Poisson noise.   Noisy or maverick pixels will contribute additional noise, but if they are relatively few in number, their contribution to the overall variance and mean, averaged over the entire array,

should be small. This technique is evaluated with the existing WHIRC PTC data in the next section.

- It is also worthwhile to carry out the difference/sum analysis for various subregions, as this could indicate variation in the conversion gain over the array. It is unlikely that such variations would be larger than the variations seen in raw flatfield images.
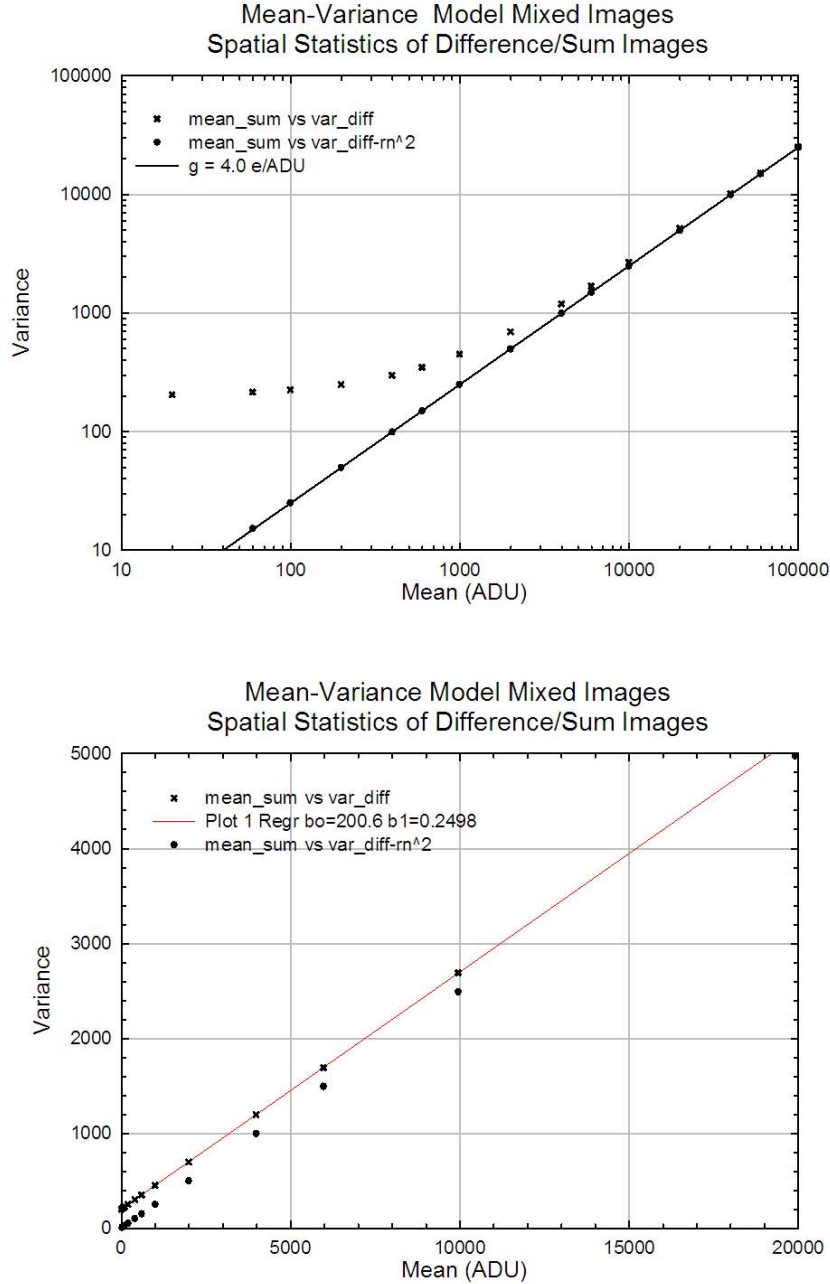

Figure 5: Mean-variance analysis using the difference/sum of two images at each intensity level, as described in Schubnell _et al_. (2006), using the mixed intensity images described in Figure 3. This method uses the spatial standard deviation of the difference image and the spatial mean of the sum image.
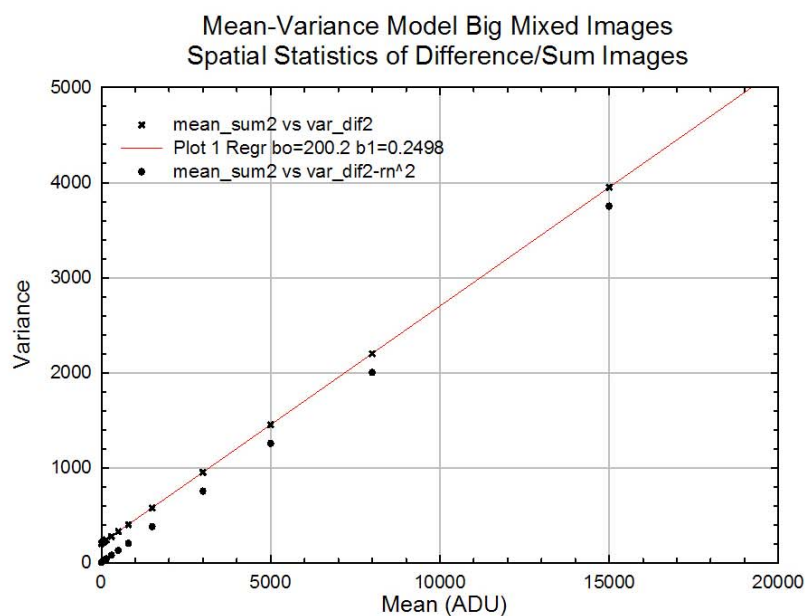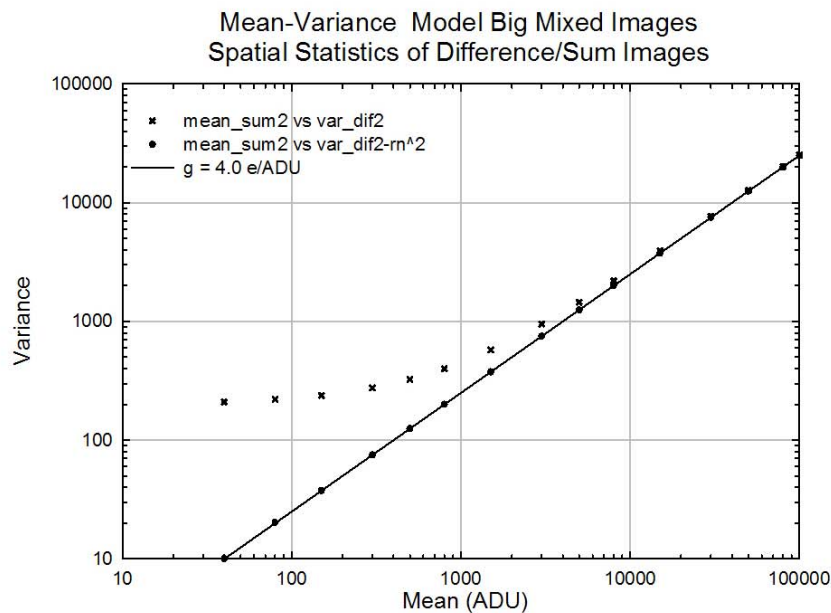
Figure 6: Same as Figure 5, except the mixed images consisted of 50% pixels at one intensity level and 50% at the next highest intensity level, as much as a factor of two greater. Even though the mean of the sum images are very different from those in Figure 5, the mean-variance analysis yields the correct gain.

4. Mean-Variance Analysis of WHIRC Detector

Several measurements of the conversion gain of the WHIRC detector were made during the instrument commissioning in concert with optimizing the performance as a function of the detector bias and readout mode. After replacing the IRAcq board in the Monsoon controller in mid-2008, we obtained a dataset on 17 September 2008 which we used for both linearity and gain (PTC) analysis. This involved illuminating the detector with a constant light source from the flatfield projector and obtaining five images each with integration times ranging from 3.8 to 85 s. To verify the constancy of the illumination for the linearity analysis, three "check" images at an integration time of 3.8 s were obtained between each increment in integration time. As noted below, while the "check" images were useful in verifying the source constancy for the linearity analysis, they complicated the mean-variance analysis using image statistics.

4.1 Image Statistics Analysis of September 2008 Data

The five images at each integration time were combined using the IRAF 'imcomb' task to generate mean and sigma images, as in section 2.1 above. The read noise was determined through a separate set of dark images to be 8.9 ADU. Image statistics were initially calculated for a number of 100×100 pixel subregions at various locations on the array, but to reduce the scatter in the statistics, we also calculated the image statistics for the entire array, excluding a 100 pixel border which seemed to have a large number of high dark current pixels. The results, plotted in Figure 7, suggest a gain of $3.8 - 4.0$ e/ADU from analysis of the logarithmic plot. These gain curves are also shown on the linear display of the PTC data.

One feature of both the logarithmic and linear plots is the tendency of the variance to rise above the linear slope (on the log plot) for large signals, suggesting an additional source of noise. Note that an unbiased fit to the linear plot would result in a different slope and an unphysical negative Y-intercept. A closer look at the individual frames showed that the mean values for large signals increased over the five data frames at a given integration time (and decreased over the three check frames immediately following; Figure 7). This is most likely a consequence of image persistence, which is known to exist at a low level in this detector from the ghost image experiment (WHIRC_IV_ghosts_080808.pdf). The drift in mean value with succeeding images will translate into an increased variance for the sigma image and yield the observed behavior.
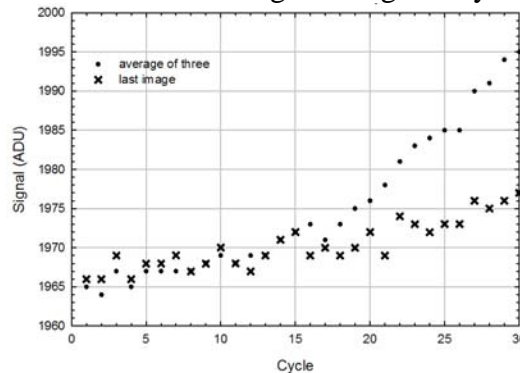

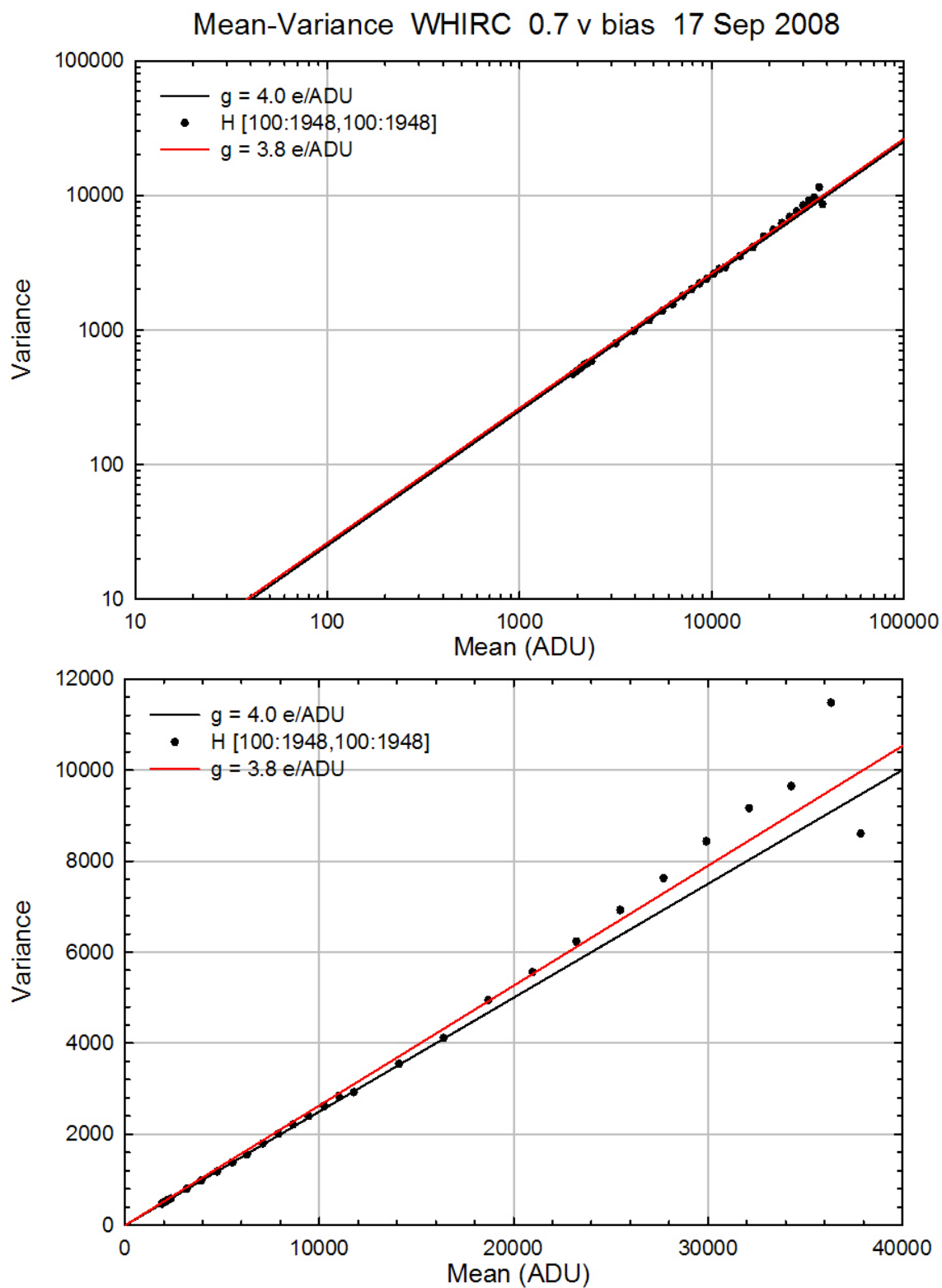Figure 7: Signal levels for the average and last image of the check images.

Figure 8: Mean-variance analysis of the PTC data of 17 September 2008 using image statistics. The logarithmic plot was used to determine a gain value 3.8 – 4.0 e/ADU for a read noise of 8.9 ADU. Curves for these parameters were overplotted on the data in the linear plot.

4.2 Sum-difference Analysis of September 2008 Data

The sum-difference analysis was applied to the same PTC dataset from 17 September 2008. Because five images were taken at each integration time, four sums and differences were calculated independently and plotted.

As with the image analysis, the logarithmic plot was used to determine the best fit to the gain (the read noise in ADU using this technique was somewhat larger, 10.4 ADU). However, the resulting gain is significantly lower, about 3.3 e/ADU vs the 3.8 e/ADU determined from the image analysis technique (Figure 9). This is an even larger difference (12%) than is seen between the two techniques for the model data (5%). The reason for this is not known. A line with the appropriate slope (0.303) and Y-intercept (220) is shown plotted on the linear representation.

One of the datasets deviates from the linear fit; this is the difference/sum of the first two images at each of the integration times where the persistence effects are likely to be most pronounced. A later PTC experiment which utilized a constant integration time and variable flux levels did not show this tendency (section 4.3).

The variance for sum image values between 14000 and 22000 ADU (i.e., single image mean values 7000 – 11000 ADU) lies significantly above the fit. Unlike the case at higher signal levels, this discrepancy is not necessarily greatest for the first two images in the set. *This effect is an artifact of the bad row (286) and column (97) in the array, which for some reason produced excess noise in the subtracted images at certain flux levels. Cleaning the images with the IRAF task 'fixpix' and the WHIRC bad pixel map eliminates the excess noise.*
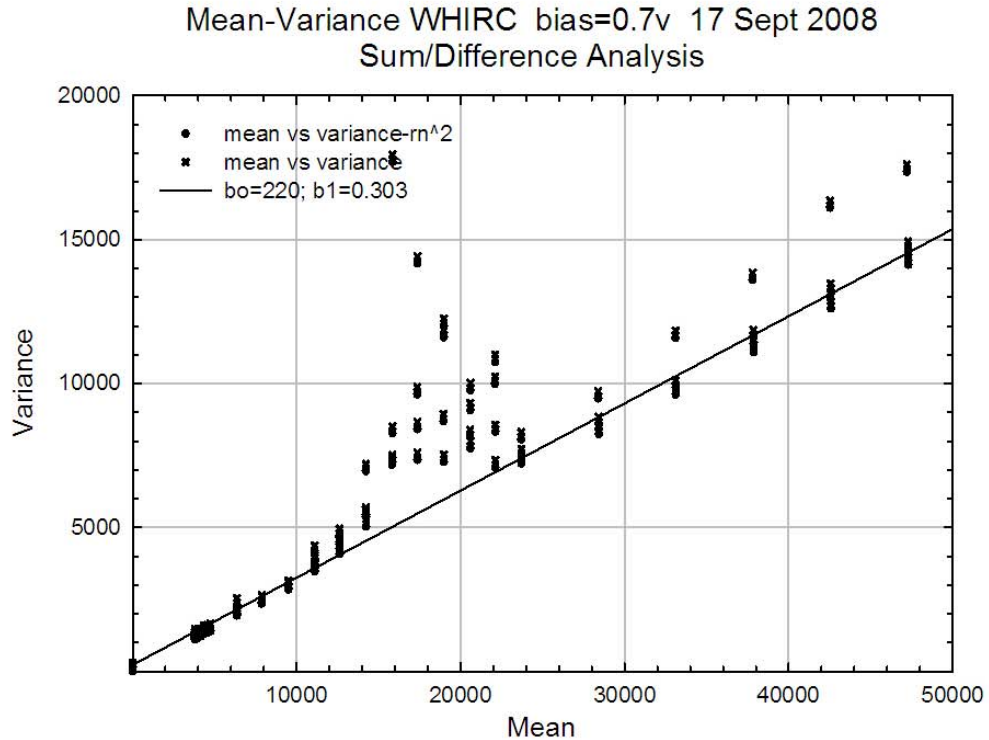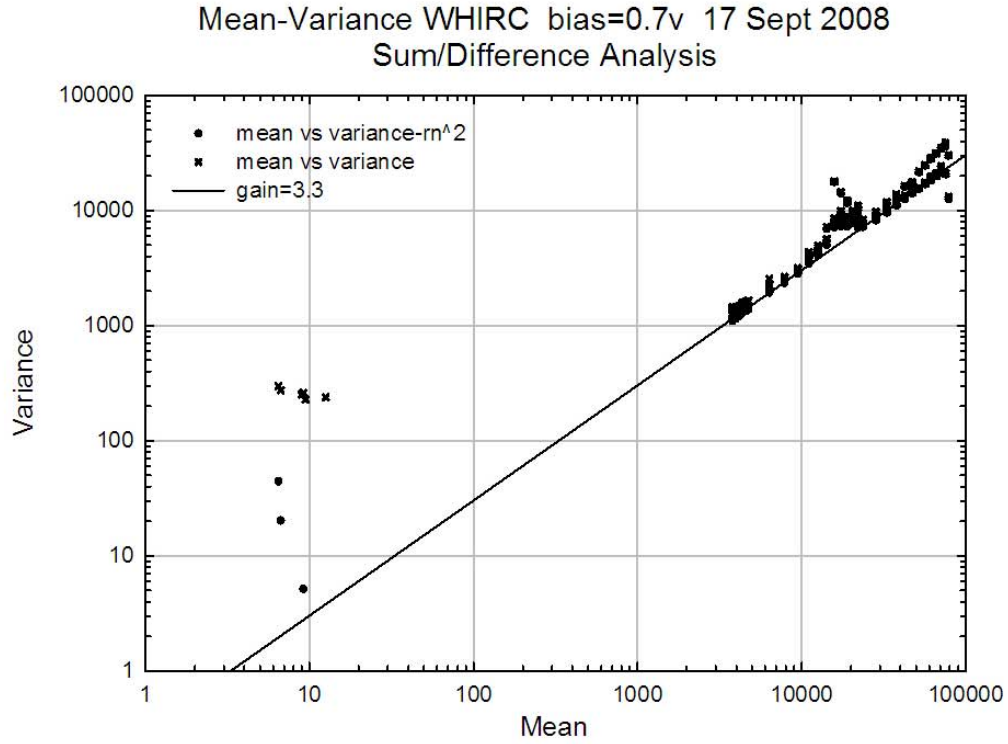
Figure 9: Mean-variance analysis of the 17 September 2008 data using the sum/difference analysis. The logarithmic plot was used to determine a gain value 3.3 e/ADU for a read noise of 10.4 ADU. Curves for these parameters were overplotted on the data in the linear plot. The deviation in the 14000 – 22000 ADU range is a result of the bad row and column, and is eliminated by filtering through the bad pixel map (see following figures).

4.3 Analysis of November 2009 and July 2010 Data

On 4 November 2009, we obtained another set of PTC data at the end of a T&E night. Since linearity was not part of the experiment, the integration time was held at a constant 5 s and the intensity of the flatfield lamp was varied. Five images were taken at each intensity setting. In addition to being quicker, this strategy also permitted us to obtain data over the entire range from no signal to saturation.

The analysis is plotted in Figure 10. Again, the logarithmic analysis was used to determine the gain and read noise (which was somewhat higher ~ 14 ADU due to some electrical pickup). The data were consistent with the gain of 3.3 e/ADU obtained from the September 2008 data, but there is less scatter among the four difference/sum images because the observing technique minimized the effects of image persistence. The results from the logarithmic analysis were overlaid on the linear plot to demonstrate the quality of the fit. In this case, anchoring the linear data plot to a Y-intercept of 400 and determining the slope would have probably independently yielded a slope close to that expected for a gain of 3.3.

The IRAF task 'fixpix' was used with the bad pixel map to clean the image. This eliminated the deviation from the fit in the 6000 – 12000 ADU range seen in the September 2008 data.

A second set of data taken 30 July 2010 and cleaned with the 'fixpix' task yields virtually identical results, with a best fit gain closer to 3.4. Since this dataset contains a greater number of points, we will adopt 3.4 as the gain value.
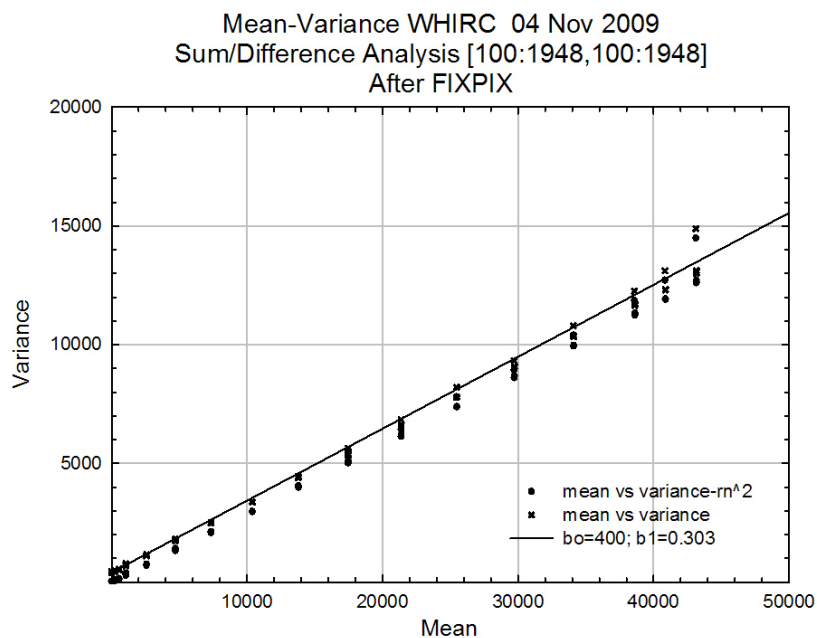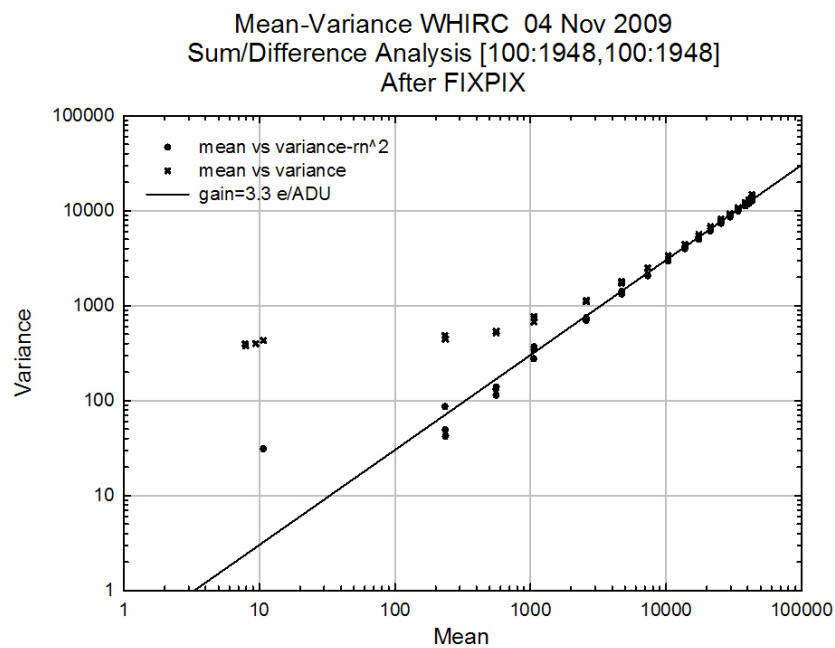
Figure 10: **Mean-variance analysis of the 4 November 2009 data using the sum/difference analysis. The logarithmic plot was used to determine a gain value 3.3 e/ADU for a read noise of 14 ADU. Curves for these parameters were overplotted on the data in the linear plot.**
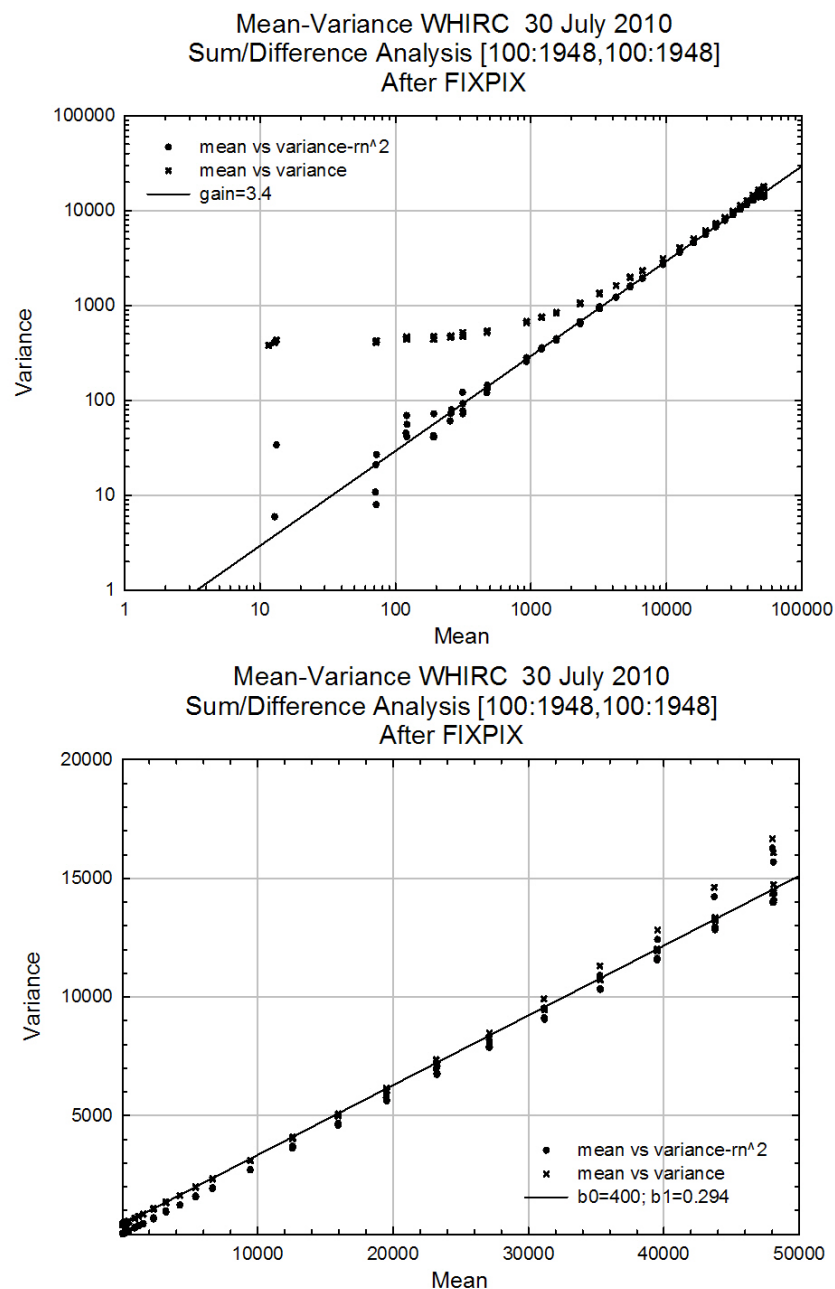
Figure 11: **Mean-variance analysis of another dataset obtained 30 July 2010 using the sum/difference analysis. The logarithmic plot was used to determine a gain value 3.4 e/ADU for a read noise of 14 ADU. Curves for these parameters were overplotted on the data in the linear plot.**