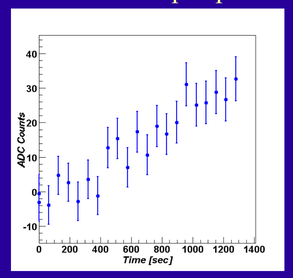
Improvements to the image processing of multiple readout HST NICMOS observations

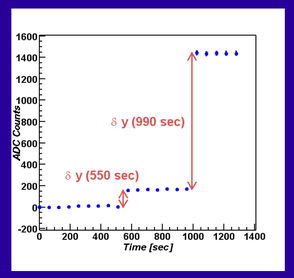
V. Fadeyev, G. Aldering, S. Perlmutter
Supernova Cosmology Project
Lawrence Berkeley National Laboratory

Our improvements include:
(1) Better CR rejection
(2) Accounting for CR inter-pixel correlations
(3) More precise error analysis

Unlike CCDs, infrared arrays can be read out non-destructively multiple times during an exposure. This allows the effect of the readout noise (R) to be reduced, and cosmic rays (CR) to be identified on a per-pixel basis.



Typical readout sequence for a pixel.



Pixel having two CR hits, at about 550 sec and 990 sec.

(1) The counts read out from a pixel are result of Poisson photonic signal P_i and readout error R :

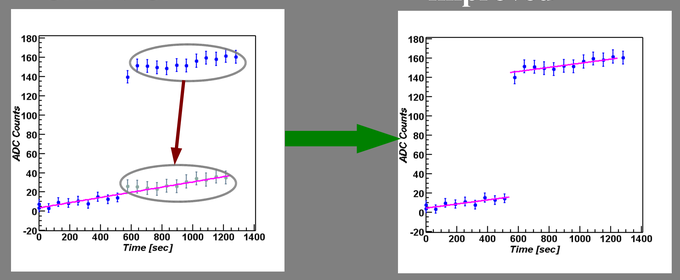
$$y_i = P_i + R \ , \qquad \sigma_i = \sqrt{P_i + R^2} \qquad (1)$$

The default data processing (CALNICA) includes a linear fit of y_i as a function of time using standard linear regression to extract the count rate.

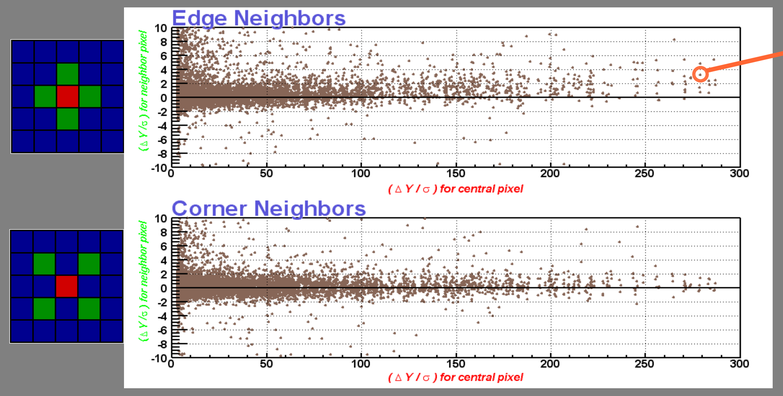
In case of a CR jump, CALNIC down-shifts the post-jump data by the two-point difference $\delta y_i = y_i - y_{i-1}$, which is not very accurate due to the measurement errors σ_i . Instead, we do a *joint fit* of data before and after the jump to linear functions with the *same slope* (b):

$$\chi^2_{2\,intervals} = \sum_{1st} \frac{(y_i - (a_1 + b*t_i))^2}{\sigma_i^2} + \sum_{2nd} \frac{(y_i - (a_2 + b*t_i))^2}{\sigma_i^2}$$

CALNICA **improved**

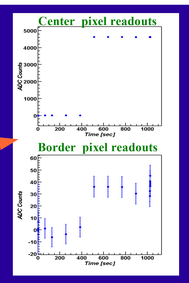


(2) Edge neighbors to the CR-affected pixels simultaneously show moderate count jumps. The corner neighbors are not affected.



Number of standard deviations for the neighbors versus the central pixel at the time of CR jump. There is a clear asymmetry for the edge neighbors.

To avoid this bias, we subject the edge neighbors to the same CR jump handling procedure.

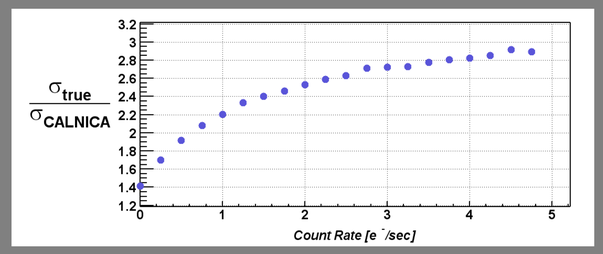


We find statistical correlation between the neighboring pixels at the time of CR hit.

(3) Applying the standard linear regression to the count values in (1) to extract the count rate is not strictly correct, since the readout values are *not independent*, due to the nature of accumulating counts.

The correct procedure would be to do the fit of correlated data with full covariance matrix: $cov(i,i+k) = P_i$, $cov(i,i) = P_i + R^2$
This is numerically intensive due to matrix inversion. According to Monte Carlo simulations, the current procedure provides an unbiased count rate estimate, which is sub-optimal only at a few percent level.

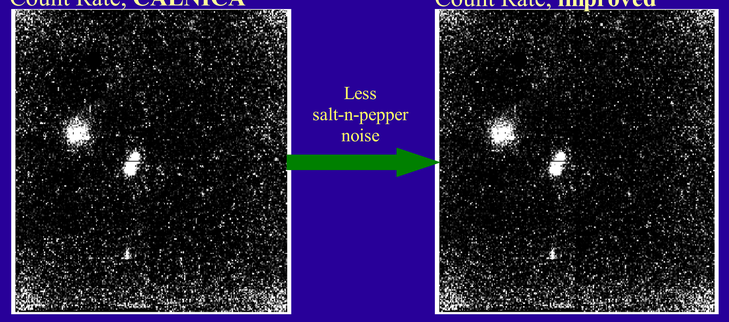
However, the CALNIC estimate of count rate error is off by a rate-dependent factor (at least 40 %).



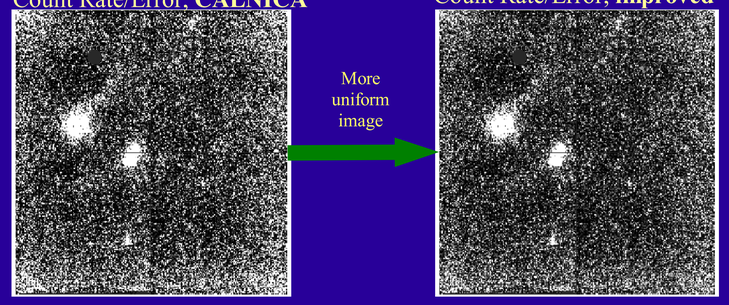
The correct procedure for the count rate error analysis requires accounting for both the readout error and the Poisson terms. One can get rid of correlations by considering independent variables $\delta P_i = P_i - P_{i-1}$ and approximating $\delta P_i \approx y_i - y_{i-1}$. This method was outlined by W.B. Sparks in NICMOS ISR 98-008 for the unweighted fit case. We derived and implemented the algorithm for our standard error-weighted fit technique.

A reprocessing example with SCP data (one exposure)

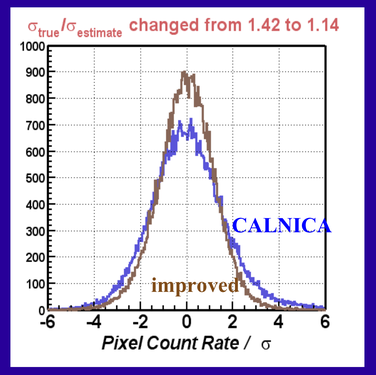
Count Rate, **CALNICA** Count Rate, **improved**

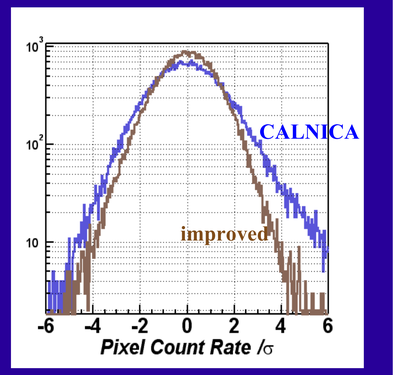


Count Rate/Error, **CALNICA** Count Rate/Error, **improved**



The improvements are somewhat difficult to see by eye, but are more obvious in the following error-normalized sky histograms.





Partial support for this work was provided by NASA through grant HST-GO-09705.01-A from the Space Telescope Science Institute, which is operated by Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555.