STScI | SPACE TELESCOPE SCIENCE INSTITUTE

WFC3 Instrument Science Report 2018-04

Persistence in the WFC3 IR Detector: An Area Dependent Model

Knox S. Long, & Sylvia M. Baggett
May 1, 2018

ABSTRACT

When the IR detector on WFC3 is exposed to a bright source or sources, the sources not only appear in the original exposure, but can appear as afterimages in later exposures, a phenomenon known as persistence. The magnitude and duration of persistence for a fixed stimulus varies somewhat across the face of the detector. Our previous attempts to characterize this variation were limited to a correction that captures only the variation in the magnitude. Here we describe a simple model which allows for variations both in the magnitude and the duration of the persistence, and then evaluate quantitatively how much improvement this model provides. We conclude that while this was a useful experiment, it does not result in a marked improvement in our ability to predict persistence in the WFC3/IR array. We discuss why this was the case, and possible paths forward.

1 Introduction

Persistence in an IR detector is the afterglow produced by light from earlier exposures. The amount of persistence in the WFC3 IR array depends on the degree of saturation in the earlier exposure, the time elapsed since the earlier exposure, and the exposure time of the earlier exposure. Persistence is higher when 1) the saturation in the earlier exposure, hereafter the stimulus image, is higher, 2) the exposure time of the stimulus image is longer, and, 3) the time delay between the stimulus image and the observer's science image is shorter.

As discussed by Long, Baggett, & MacKenty (2015a), the area-averaged persistence of the WFC3 IR detector can be described in terms of a model of the form:

$$P = A \left(\frac{t}{1000\ s} \right)^{-\gamma} \quad (1)$$

where A and γ are functions of fluence total counts and exposure time. This model is currently being used to produce data products available through MAST that estimate persistence in all WFC3/IR observations and its parameters were determined from the 8 visits of Program 13572. Each visit started with a single external image of a globular cluster (either 47 Tuc or Omega Cen), the stimulus image, followed by a long series of darks within which persistence could be measured. The exposure times in the stimulus images range from 49 to 1403 s. Interpolation is used to estimate persistence for exposure times not explicitly used in the persistence calibration data. The full persistence model includes a "correction flat" described by Long, Baggett, & MacKenty (2015a), which provides a zeroth-order correction for changes in persistence amplitude across the face of the detector. This model does not address location-dependent variations in the persistence decay rate with time (γ in Eq. 1 above).

Here we describe a straightforward extension to the model currently in use that captures at least partially the variation in γ across the face of the detector.

2 A Simple Modification of the Existing Persistence Model

To first order, persistence in the WFC3/IR array decays with a power law with an amplitude that varies as a function of position, the length of the exposure and the fluence of the image causing persistence. With sufficient data, it would in principle be possible to determine the amplitude A and power law slope (γ) for each pixel of the detector which would obviate the need for a correction flat. Many properties of the IR detector, including the gain and linearity are calculated on a per pixel basis. However, this is not possible for persistence as little information was obtained about persistence in the flight device prior to launch. Performing such a calibration on-orbit is prohibitively expensive in terms of observing time. Fortunately, persistence appears to vary slowly over the face of the detector (see Long, Baggett, & MacKenty (2015b)) and as a result one can consider various alternatives to account for the variations.

The simplest of these alternatives, or the easiest to implement at least, is to divide the detector into N x M regions, determine A and γ for each region, and use them to calculate the persistence. In principle, one could create a model for each pixel this way, e.g. by using a 100 x 100 box around each pixel to calculate A and γ, but a simpler approach is to use some kind of "correction flat" to smooth over the edges. The advantage is that for a given set of calibration data, it is relatively straightforward to explore trade-offs involving the number of regions.

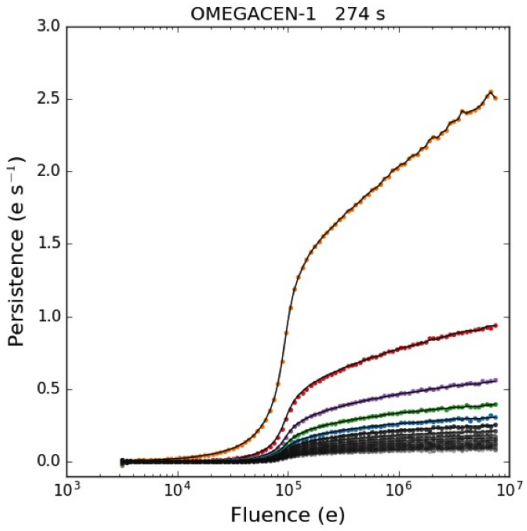

Figure 1. The persistence following visit 1 of program 12351, which consisted of a single 274 second exposure of the globular cluster Omega Cen followed by a series of darks. Each curve represents the persistence measured in one of the darks. Persistence is highest in the first dark (top curve) taken immediately after the stimulus image and decays to progressively lower levels in each successive dark (lower curves).

3 Data and its preparation

The data used in this analysis consists of a selection of the WFC3 IR observations from Cycles 18 through Cycle 23 obtained explicitly for studying persistence. Each of these visits started with a single external exposure of a globular cluster followed by a long series of darks. We only consider cases where there are at least four visits that are similar in the sense that the stimulus image is taken with the same exposure time. Our motivation for choosing this criterion is to assure that we have the most possible data for each stimulus exposure time we analyze.

Table 1. Observation Log

ProgID	Visit	Dataset	Obs Date	Target	Filter	Exp. (s)	Saturated (%)
14381	91	id1s91krq	2016-04-05	OMEGACEN-3	F140W	149	4.3
14381	81	id1s81whq	2016-03-31	OMEGACEN-2	F140W	149	4.4
14381	71	id1s71o5q	2016-03-23	OMEGACEN-1	F140W	149	4.6
13572	02	icgk02zpq	2014-03-06	NGC-104	F125W	149	11.9
14015	02	icrq02kqq	2015-03-24	OMEGACEN-1	F140W	149	4.7
14381	01	id1s01sqq	2016-01-07	OMEGACEN-1	F105W	274	7.2
12351	01	ibmf01vfq	2011-01-27	OMEGACEN-1	F110W	274	13.5
12351	03	ibmf03nbq	2011-02-01	OMEGACEN-3	F110W	274	12.7
12351	AC	ibmfacnyq	2011-05-25	OMEGACEN-2	F110W	274	13.2
14381	F1	id1sf1xyq	2016-06-10	OMEGACEN-3	F125W	499	14.8
14381	D1	id1sd1afq	2016-05-09	OMEGACEN-1	F125W	499	14.9
14381	E1	id1se1brq	2016-06-03	OMEGACEN-2	F125W	499	14.7
13572	03	icgk03jiq	2014-02-04	OMEGACEN-1	F110W	499	25.9
14381	31	id1s31ppq	2016-02-02	OMEGACEN-3	F125W	599	17.1
14015	05	icrq05enq	2015-03-28	OMEGACEN-1	F125W	599	18.6
14381	11	id1s11jiq	2016-01-12	OMEGACEN-1	F125W	599	18.5
14381	21	id1s21j0q	2016-01-19	OMEGACEN-2	F125W	599	18.4
14381	A1	id1sa1haq	2016-04-15	OMEGACEN-1	F127M	899	5.1
14381	B1	id1sb1w6q	2016-04-29	OMEGACEN-2	F127M	899	5.1
14015	06	icrq06xkq	2015-04-24	OMEGACEN-1	F125W	899	28.4
14381	C1	id1sc1dfq	2016-05-20	OMEGACEN-3	F127M	899	5.0
14381	41	id1s41dkq	2016-02-12	OMEGACEN-1	F127M	1199	7.3
14381	51	id1s51csq	2016-03-07	OMEGACEN-2	F127M	1199	7.2
14381	61	id1s61ohq	2016-03-16	OMEGACEN-3	F127M	1199	6.8
14015	07	icrq07d8q	2015-04-25	OMEGACEN-1	F127M	1199	7.3

The observations are listed in Table 1 ordered by exposure time. The table lists the (1) program number, (2) visit number[1], (3) dataset, (4) observation date, (5) target name, (6) filter, (7) exposure time and (8) percentage of pixels saturated in the external exposure. There are 6 different stimulus exposure times ranging from 149 s to 1199 s in the 25 different visits. There are multiple pointings within Omega Cen, denoted by the '1', '2', and '3' in the target name field.

A significant caveat is that even for visits with similar stimulus images, some differences remain. The most obvious disparity is that in several cases where the stimulus image has the same exposure time, the filters are different and as a consequence the number of pixels that are exposed to saturated flux levels differs. Even when the filter and exposure time match for the stimulus images, individual pixels are still illuminated to different levels in each visit due to small offsets in pointing position and field of view orientation. Persistence varies on large spatial scales with the highest and lowest persistence in the upper left and lower right quadrants of the detector, respectively (Long, Baggett, & MacKenty 2015b). Because stars of a given brightness are not distributed precisely uniformly across the file, pointing offsets will inevitably induce some apparent variation in persistence between different visits.

As noted earlier, each visit in this program consists of an external stimulus exposure followed by a series of darks. Bright stars in the external exposure produce pixels at a variety of saturation levels across the face of the detector. Note that in creating a model for the average persistence, we are explicitly assuming the number of bright stars is fairly uniform across the face of the detector. As a preliminary step for the data analysis described here, we have reprocessed all of the data with version 3.1.6 of CALWF3. To have the calibration software apply the gain correction to the darks, we activate the flatfield correction and use a "unity flat" (containing only 1's). We then construct persistence curves for each of the visits using a standard set of stimulus levels for each of the darks (after subtracting our best estimate of the dark level in each dark exposure). In creating these curves, we subtract a residual dark current, using the median value of pixels with a stimulus in a certain range, usually pixels exposed to a stimulus less than 10,000 e. An example of these persistence curves for one of the visits is presented in Fig. 1.

[1]For scheduling reasons, each persistence observing sequence in program 14381 was split into visit pairs: the first visit contains the stimulus image plus the first few post-stimulus darks followed immediately by the second visit containing the remaining darks. In Table 1 we list only the first visit of a pair i.e. the visit in which the external stimulus exposure was acquired.

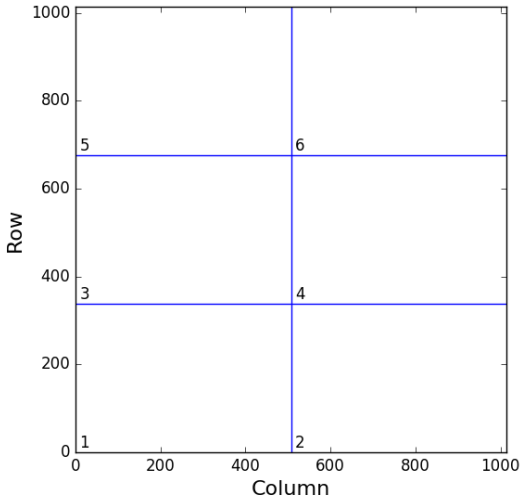

Figure 2. An illustration of our numbering scheme for dividing the array into 6 separate regions, where N is 2 and M is 3. Note that for a 2 x 2 grid this naming convention results in subsection numbers that do not match the nominal WFC3 IR quadrant nomenclature (1-4 counter-clockwise starting at the upper left).

4 Analysis Approach

4.1 A power law model for NxM regions

To test the model, we divide the IR detector into N x M regions, where N is the number of sections along the row axis and M is the number of sections along the column axis of the detector. Thus a 2x3 model divides the detector into 2 horizontal regions and into 3 vertical regions, numbered according to Fig. 2. A 1x1 model corresponds to the model described by Long, Baggett, & MacKenty (2015a). We assume that each region behaves according to Eq. 1, but independent values of A and γ are calculated for each region.

Operationally, for each region of the detector and using a standard grid of stimulus values, the first step in the fitting process is to compute curves like the one shown in Fig. 1 for all visits with a particular stimulus exposure time. We estimate the persistence at a particular stimulus level from the median value of the pixels that have good data quality flags in both the stimulus and the dark image. We assign an error to the fit given by the standard deviation of values in the interval divided by the square root of the number of data points in the interval.

We then determine the best fit for all of the visits with a given stimulus level and exposure time for each region to produce optimum fit values of A and γ. We cannot (usually) combine data from multiple visits before fitting because the delay times are not always the same, even when the stimulus exposure time was the same.

The results of such a fit for a 4 x 4 subsection grid based on the set of calibration data with a stimulus exposure time of 599 s are shown in Fig. 3. In a general sense all of the curves look rather similar. The amplitude of persistence (at 1000 s) rises rapidly as the fluence reaches saturation (70,000 e) and rises relatively slowly after that. The power law indices are steep at low stimulus levels and flatten to a value close to 1 above saturation. Not only is the amplitude of persistence at 1000 s higher at higher stimulus levels, it lasts longer. The amplitude of the fitted persistence is higher than average in the upper left quadrant of the detector, especially cells 9 and 13, as expected from previous analyses of the spatial persistence variations (Long, Baggett, & MacKenty (2015b)). Furthermore, at a given stimulus level the power law slope in the upper left quadrant of the detector is less than average, indicating persistence lasts longer there as well.

The general trends in each of the regions are also similar to that reported in Long, Baggett, & MacKenty (2015b) for what was effectively a 1 x 1 model. The amplitudes all have similar characteristic shapes with A increasing with exposure time. The power law slope is steeper at low stimulus values and flattens to a value of about 1 as the stimulus approaches saturation. Persistence is thought to be due to a trapping phenomenon, as discussed by Smith et al (2008a,2008b). Except for the fact that one typically expects exponential rather than power law decays in such circumstances, the other characteristics of the persistence do arise naturally in this paradigm.

Several other points are worth noting. The individual 4 x 4 fits in Fig. 3 show evidence of increased noise at very low and very high fluence levels. At high levels, this is likely due to a lack of pixels with the appropriate stimulus level, and thus one expects the results to be noisy. To pursue this study further, one option is to consider binning of the stimulus levels. At low fluence levels, the limiting issue is likely the time variable dark current[2] and the fact that we are trying to measure a persistence signal that is very close to this background.

4.2 A correction flat

Within any region, the persistence can still vary spatially. To account for this effect, we use the same procedure as was applied previously. We assume that the actual persistence can be represented with a correction R(x,y) to the model M(t) such that the actual persistence as a function of position and time is given by

$$P(x, y, t) = (1 + R(x, y))M(t) \tag{2}$$

where M is the NxM power law model and all of the residual area dependence is captured by R(x,y). Effectively this means we are assuming that we have the correct average persistence when R is being calculated.

If that is the case, then

[2]Variations in the dark rate can be significant: the mean rate is about 0.05 e s^{-1} and the standard deviation about the mean is ± 0.03 e s^{-1} (Hilbert & Petro 2012; Sunnquist, Baggett, & Long 2017).

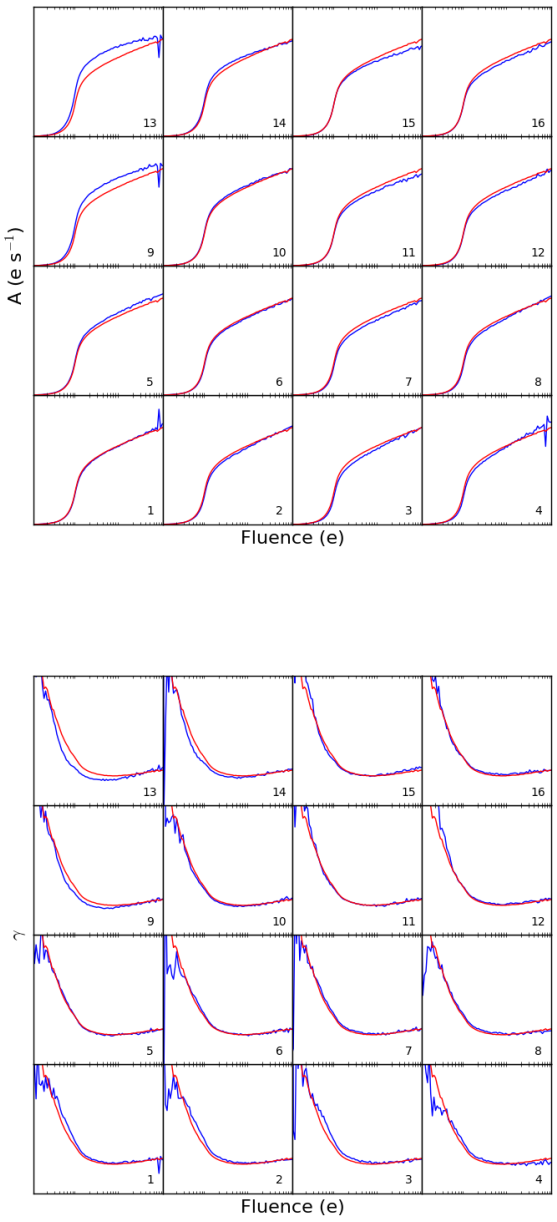

Figure 3. Fits for a 4 x 4 model using all of the calibration data for a stimulus image with exposure time of 599 s. The top panel shows the fitted amplitudes to each section of the detector in blue, and the average of the fitted amplitudes in red. The bottom panel shows the fitted power law indices. The fluences on the x axes are plotted logarithmically from 10,000 to 10^7 e. The amplitudes on the y axes of the top panel are plotted linearly from 0 to 0.7 e s^{-1}. The power law indices in the bottom panel are plotted linearly from 0 to 1.7.

$$E(x,y) = P(x,y) - M(t) = R(x,y)M(t) \tag{3}$$

where E(x,y) is the error in a persistence-subtracted dark, which, by assumption, is time-independent. Since M is a simple number independent of position (except implicitly since M is dependent on the fluence in the earlier exposure), and we know that by construction the average of R(x,y) is about zero, i.e. $< R(x,y) > \sim 0$, then we argue that any departure from 0 in E(x,y) is due to our lack of knowledge of the dark current. The $\sim$ in this equation arises from the fact that M is not really a constant.

Therefore for a single image

$$R(xy) = \frac{E(x,y) - < E(x,y) >}{M} \tag{4}$$

In order to calculate R, we use a series of visits comprised of Tungsten flat exposure for the stimulus image followed by a series of darks in which the persistence is measured. Each dark provides a measure of R(x,y) and therefore one must decide how to weight the exposures. Giving all of the exposures identical weight would be a mistake, as the noise in the ratio is higher when there is very little persistence. Therefore, we decided to weight by the mean persistence in the entire image, that is

$$R(x,y) = \frac{\Sigma_i(< M_i > \frac{E_i(x,y) - < E_i(x,y) >}{M_i})}{\Sigma_i < M_i >}, \tag{5}$$

where $< M_i >$, refers, as is our standard practice, to the median (rather than the average) of the model for persistence in the ith dark. The correction flat then is given by $1 + R(x,y)$.

For the purposes of this report, we used visits from programs 12089 and 12551 that begin with a tungsten lamp flat. These are listed in Table 2. Any visits preceded by exposures which would have caused a significant amount of persistence were eliminated from the analysis. The columns of the table are (1) the program ID, (2) the visit, (3) the dataset name of the Tungsten stimulus exposure, (4) the observation date, (5) the filter used for the Tungsten exposure, (6) the Tungsten exposure time, (7) the median fluence value in the Tungsten exposure, and (8) the fraction of the pixels in the image which were saturated. The visits we constitute a heterogeneous set, with stimulus exposure times ranging from 32 to 2602 s, and exposure levels (based on the fluence) ranging from about 1/3 to 20 times saturation. The data were re-reduced using the same procedures as described in Sec. 3.

From a practical perspective, the process for generating the correction flat is straight forward: Generate a model without the correction flat using the procedure described in Sec. 4.1. Use the model without a correction flat to remove the average persistence from a set of Tungsten visits. Construct the correction flat $1 + R(x,y)$ as described here. Smooth as appropriate.

The correction flat for the 4x4 model described earlier, smoothed with a 5x5 median filter, is shown in Fig. 4. This flat is fairly typical. The boundaries of the 4x4

Table 2. Tungsten Visits

ProgID	Visit	Dataset	Obs Date	Filter	Exp. (s)	Median Stimulus (e)	Saturated (%)
12089	05	ibel05s2q	2010-05-23	F105W	32	23126	0.0
12351	21	ibmf21enq	2011-04-08	F105W	72	52325	0.0
12089	02	ibel02g7q	2010-05-21	F105W	72	52489	0.0
12351	11	ibmf11qbq	2011-05-04	F105W	72	52225	0.0
12351	A1	ibmfa1rkq	2012-05-19	F105W	72	51966	0.0
12351	63	ibmf63jeq	2011-05-15	F125W	72	82967	95.5
12351	66	ibmf66klq	2012-03-13	F160W	72	82339	92.8
12351	25	ibmf25kmq	2011-04-24	F105W	102	74138	78.4
12351	15	ibmf15a3q	2011-05-06	F105W	102	73185	73.2
12351	A5	ibmfa5kiq	2012-05-26	F105W	102	73293	73.9
12351	7Z	ibmf7zbrq	2012-05-11	F160W	122	139900	99.6
12351	A6	ibmfa6h3q	2012-05-29	F105W	152	108912	98.8
12351	26	ibmf26hlq	2011-04-23	F105W	152	110572	98.9
12351	16	ibmf16omq	2011-04-06	F105W	152	110386	98.9
12351	A2	ibmfa2ywq	2012-05-24	F105W	202	144619	99.5
12351	12	ibmf12ytq	2011-04-29	F105W	202	143083	99.5
12089	01	ibel01p1q	2010-05-09	F105W	202	144081	99.5
12351	61	ibmf61wbq	2011-06-16	F098M	227	69822	47.9
12351	81	ibmf81juq	2012-05-29	F098M	227	71038	60.7
12351	AD	ibmfadeyq	2012-07-02	F153M	252	71825	59.9
12351	82	ibmf82zjq	2012-05-31	F127M	252	72090	62.9
12351	62	ibmf62vbq	2011-04-07	F127M	277	79176	89.0
12351	65	ibmf65g9q	2011-04-09	F153M	277	79222	85.8
12351	A7	ibmfa7m1q	2012-05-26	F105W	302	215755	99.8
12351	17	ibmf17b4q	2011-05-14	F105W	302	215072	99.8
12351	27	ibmf27xcq	2011-04-21	F105W	302	216617	99.8
12351	84	ibmf84dvq	2012-06-04	F098M	452	138944	99.4
12351	41	ibmf41lrq	2010-12-21	F140W	499	821361	99.5
12351	85	ibmf85gvq	2012-07-02	F127M	502	142680	99.5
12351	71	ibmf71qqq	2012-05-09	F153M	502	142660	99.6
12351	68	ibmf68aqq	2012-05-04	F127M	502	141943	99.5
12351	86	ibmf86nfq	2012-07-28	F153M	502	142471	99.6
12351	18	ibmf18neq	2011-04-06	F105W	502	358838	99.9
12351	A8	ibmfa8wtq	2012-05-20	F105W	502	356476	99.9
12089	03	ibel03e2q	2010-05-15	F105W	652	469004	99.9
12351	23	ibmf23hsq	2011-04-18	F105W	652	465823	99.9
12351	89	ibmf89ras	2012-08-01	F164N	1002	70141	50.8
12351	AB	ibmfabyxq	2011-03-18	F105W	1002	716061	100.0
12351	A9	ibmfa9siq	2012-05-09	F105W	1002	718009	99.9
12351	88	ibmf88hmq	2012-07-14	F132N	1102	70527	53.6
12351	87	ibmf87p7q	2012-06-13	F126N	1302	69561	46.4
12351	70	ibmf70ddq	2012-03-16	F164N	2002	137204	99.4
12351	92	ibmf92a9q	2012-06-18	F164N	2002	138099	99.4
12089	04	ibel04l3q	2010-05-11	F105W	2002	1432890	100.0
12351	AA	ibmfaahzq	2011-03-20	F105W	2002	1424603	100.0
12351	24	ibmf24neq	2011-04-02	F105W	2002	1435165	99.9
12351	A4	ibmfa4c1q	2012-04-30	F105W	2002	1406198	99.9
12351	AE	ibmfaelxq	2012-07-03	F132N	2402	153838	99.3
12351	90	ibmf90ieq	2012-05-29	F126N	2602	139284	99.1

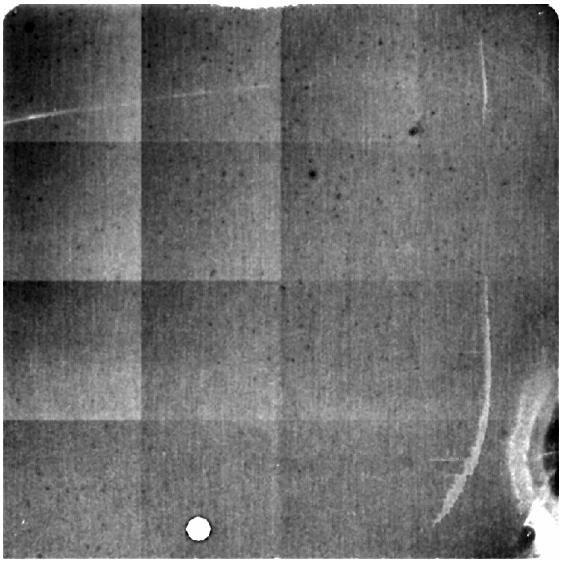

Figure 4. The correction flat for the 4x4 model, linearly scaled from 0.8 (light) to 1.2 (dark).

subsections are apparent. The power law model removes the average persistence across each region but variations across individual regions are apparent. Persistence generally increases towards the upper left of the WFC3 IR array, resulting in higher values of the correction flat in the upper left of many of the individual cells. Various image defects are also visible, such as the circular area of dead pixels near the bottom left of center, the arc-shaped area of pixels with lower-than-average quantum efficiency at lower right. Many of the pixels in defect regions were flagged out when the power law fits were made so the model over-predicts persistence in these areas. The correction flat helps to correct for this. There are a few well-defined regions in the upper right quadrant of the detector that have anomalously high persistence; these had been seen previously (Long, Baggett, & MacKenty (2015b)).

5 Tests of the Persistence Model

Given a persistence model, one would like to know how well it actually removes persistence. For exactly the same reasons that we cannot construct a persistence models on a pixel by pixel basis, we cannot measure the accuracy with which persistence is removed on a pixel by pixel basis. As a result, our approach has been to measure the error in persistence as a median of the residual persistence in N'xM' regions of the detector, where usually N' and M' represent a different (finer) scale than that used to created a model.

The quality of the subtraction is indicated in Fig. 5 and 6, where we show the error in the persistence subtraction for the 4 x 4 model for pixels that had a stimulus

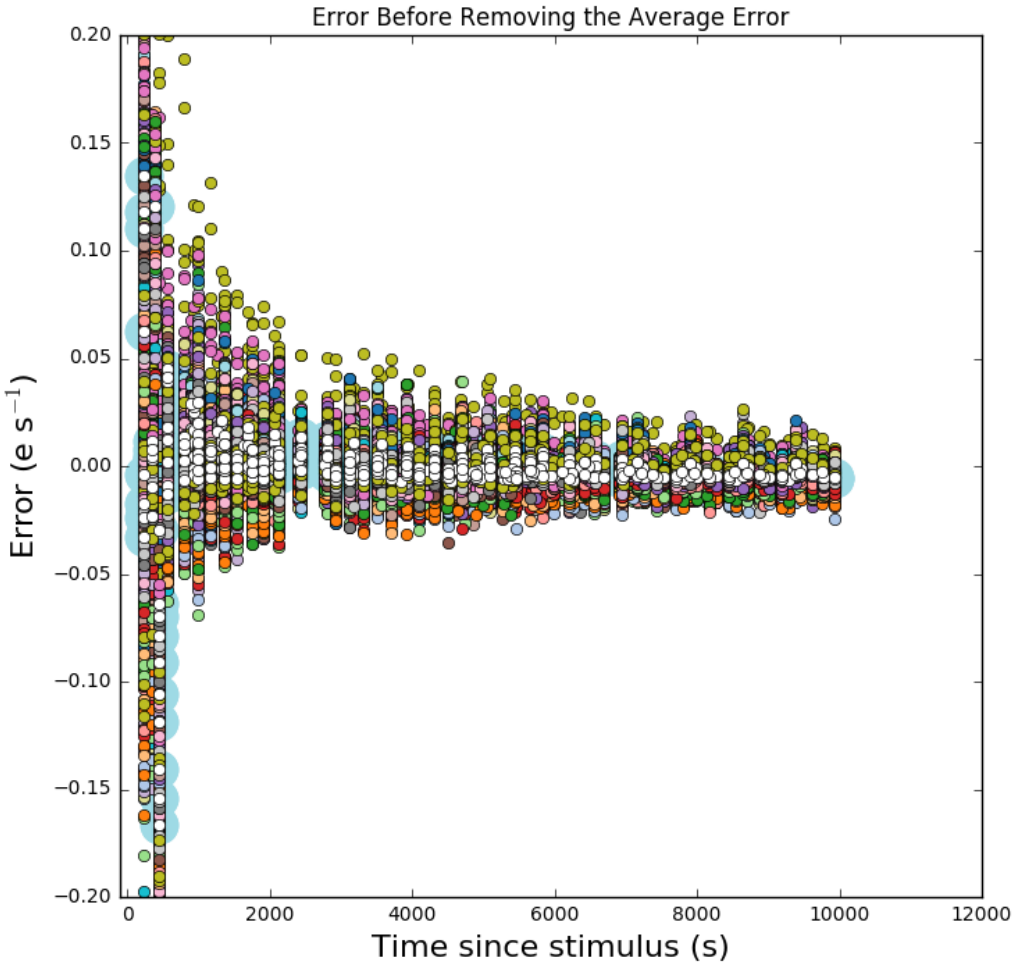

Figure 5. Errors (observed - model) in the persistence subtraction of the 4x4 persistence model as seen in each of 8 x 8 sections of the detector as a function of the time since stimulus. The white circles show the average error over the entire array; the circles with other colors show the error in the individual cells.

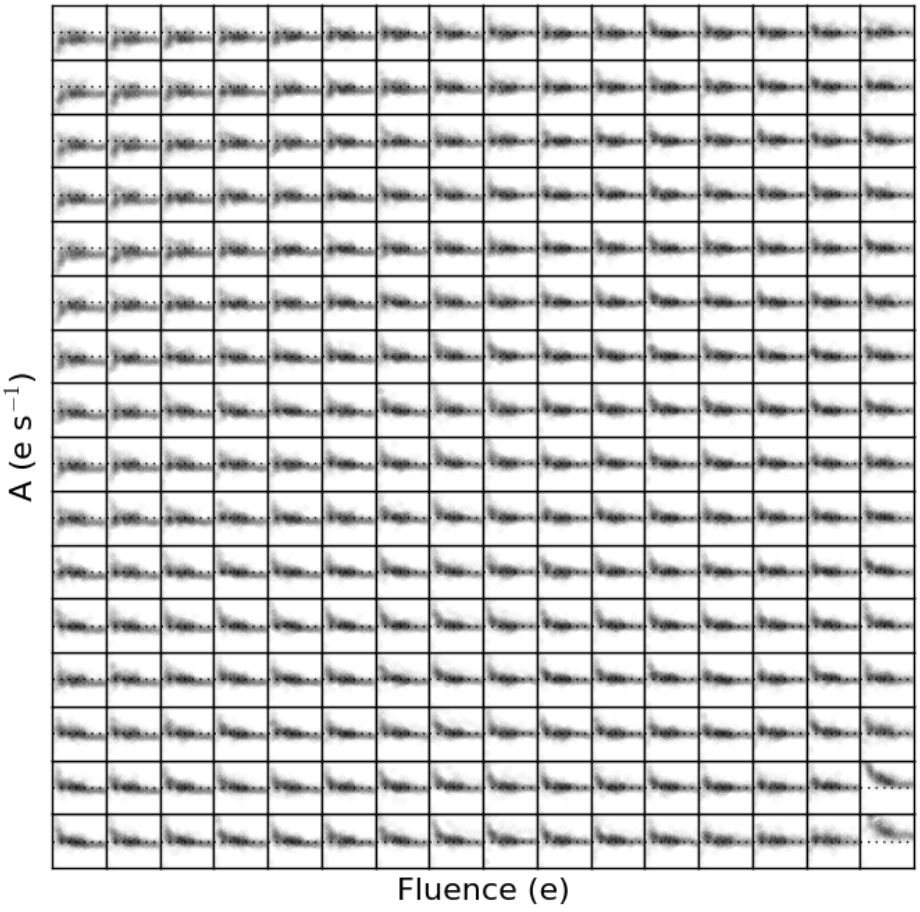

Figure 6. Errors (observed - model) in the persistence subtraction of the 4x4 persistence model as seen in each of 8 x 8 sections of the detector in the form of a mosaic. The vertical scale for each of the subpanels ranges from -0.05 to 0.05 e s^{-1}.

Table 3. Errors - 4 x 4 Model

Stimulus Exposure	Mean Error (0-1500s)	σ (0-1500s)	Ratio (0-1500s)	Mean Error (>1500s)	σ (>1500s)	Ratio (>1500s)
149	-0.006	0.076	0.078	0.004	0.008	0.116
274	-0.005	0.060	0.049	0.002	0.007	0.067
499	-0.002	0.063	0.071	0.000	0.006	0.064
599	0.005	0.052	0.063	-0.003	0.006	0.071
899	0.001	0.028	0.032	-0.002	0.007	0.067
1199	0.003	0.029	0.029	-0.003	0.006	0.065

of between 10^5 and 5×10^6 e when applied to the datasets listed in Table 1. These are the same datasets used to create the calibration which is not the ideal way to test the accuracy of the model, but there is no other set of data of comparable size. As shown in Fig. 5, the error in persistence in each of the 64 cells is less than about 0.03 e s^{-1} within 1500 s of the stimulus in essentially all cases. To the extent that there are systematic errors in the model there seems to be a slight tendency to under-subtract the persistence at early times, with the largest errors near the lower right corner of the detector where the most defects exist in the detector (lower right in Fig. 6).

Similar estimates of the errors for a 1x1 model are shown in Fig. 7. Qualitatively, the results are very similar to the results for the 4 x 4 model. The straightforward interpretation, since the correction flat is perfectly capable of removing amplitude variations, is that differences in the power law slope do not contribute greatly to the overall model. Indeed, the one place where there is a significant difference in the residuals is in the upper left corner of the detector, where as indicated in Fig. 3 the power law dependence deviates significantly from the average.

One can make some of these statements more quantitative. Table 3 lists (1) the stimulus exposure time for which the errors are calculated, (2) the mean error in the persistence subtraction 0 to 1500 s after the stimulus exposure, (3) the standard deviation of the errors calculated for the individual cells, (4) the fraction of the persistence which remains (5) the mean error in the persistence subtraction for delay times greater than 1500 s, (6) the standard deviation of the errors for delay times greater than 1500 s, (7) the fraction of persistence which remains for times greater than 1500 s. Except for the stimulus exposure time, all of the values in the table are in e s^{-1}. The pixels that were used to generate these estimates had a fluence in the stimulus image of 10^5 and 5×10^6 e. To clarify, the error in a single cell is the median of the error for all pixels in a cell which had a fluence in the stimulus image between 10^5 and 5×10^6 e. The mean in the table is the average of the 64 cell measurements and the standard deviation is the standard deviation of these 64 measurements. The ratio is the ratio of the absolute value of the residual persistence to the model prediction; it is a reasonable representation of the fraction of the persistence that remains.

The values for the mean error and the standard deviation should be compared to

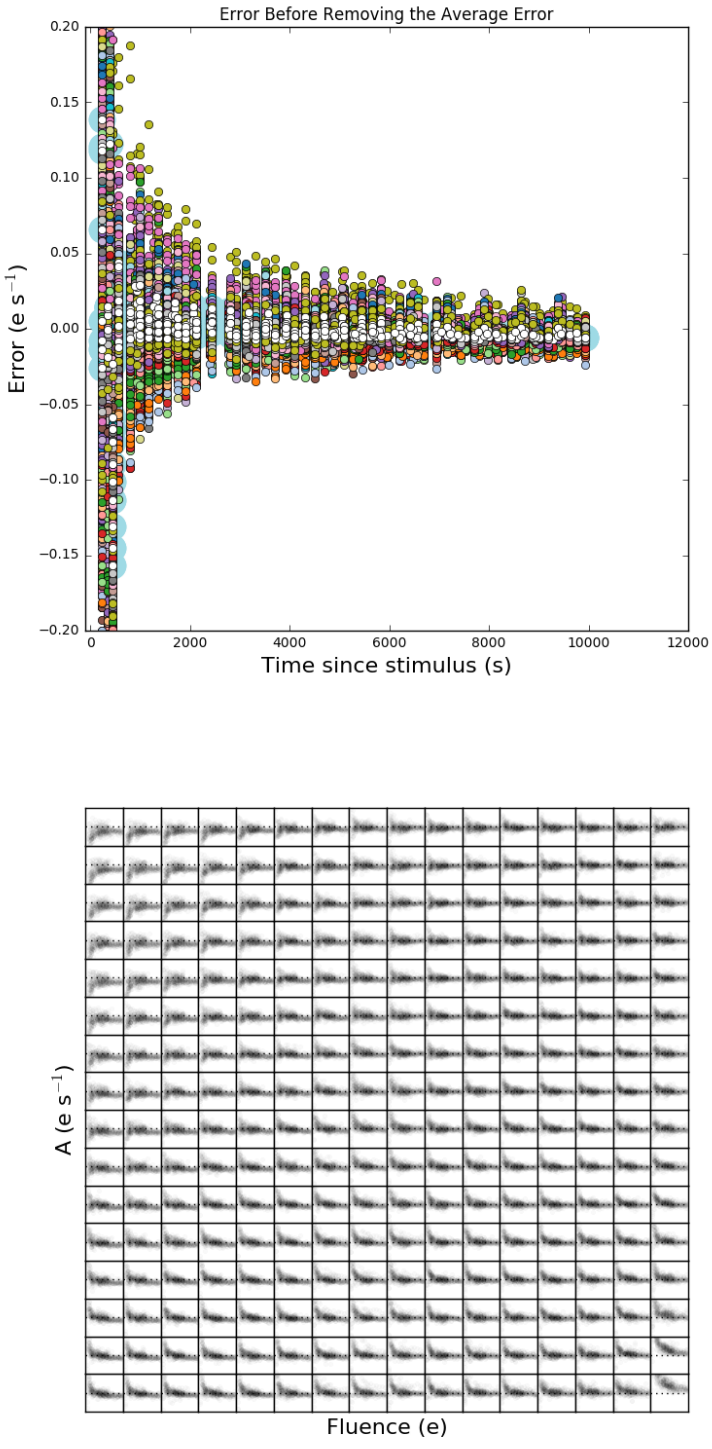

Figure 7. Identical to Fig. 5 and 6, except for a 1 x 1 model.

Table 4. Errors - 1 x 1 Model

Stimulus Exposure	Mean Error (0-1500s)	σ (0-1500s)	Ratio (0-1500s)	Mean Error (>1500s)	σ (>1500s)	Ratio (>1500s)
149	-0.008	0.084	0.084	0.003	0.009	0.117
274	-0.003	0.064	0.050	0.001	0.007	0.062
499	-0.003	0.071	0.073	-0.000	0.006	0.063
599	0.004	0.060	0.068	-0.004	0.006	0.075
899	0.001	0.053	0.048	-0.003	0.007	0.071
1199	0.003	0.055	0.049	-0.004	0.007	0.071

the dark current which is about 0.05 e s^{-1}. The model on average does very well in estimating the persistence at both short and long time delays. At short times, about 95% of the persistence is removed, but the standard deviation is of order the dark current. Interestingly the standard deviation is worse at short stimulus exposure times than at long stimulus exposure times. Finally, at long delay times (greater than 1500 s), both the means and the standard deviations are well below (typically less than about 10%) of the dark current, but the fraction of persistence remaining tends to be somewhat higher.

The same information for the 1x1 model is provided in Table 4. The results are very similar to those for the 4x4 model, perhaps marginally worse, but clearly not enough to make a difference in any practical sense.

We have experimented with other choices for dividing the detector into subregions to create models, e.g 2x2 or 3x3. The quality of the predictions of these models is very similar to the ones presented here. At higher values, e.g. 8x8, the fits become worse; this is surely due to the lack of sufficient calibration data. Our overall conclusion is that while this was a useful experiment, the improvement in performance using an area-dependent model of this type does not merit implementation as part of the persistence pipeline.

6 Discussion

As described above, our attempt to improve our model of persistence using an N X M model was not very successful. Why is this case? There are three basic possibilities: (1) Persistence varies in nominally identical data sets, and this places a limit on how well we can measure persistence. (2) Our ability to carry out the fits is compromised in some fashion. The statistical errors could be larger than we expect and/or we are not accurately subtracting the dark current. (3) The underlying assumptions of the model are incorrect. That is, there may be departures from power law that affect the results and/or using a spatially-varying amplitude correction flat introduces errors. In this section, we investigate each of these possibilities in detail in an attempt to find a path toward a better model of persistence.

6.1 Variations in the underlying Persistence

For the models we have discussed thus far, we combined calibration data from multiple visits with the same stimulus exposures as using more data should allow us to measure A and γ more accurately in the subregions of the detector. We were aware at the time that occasionally there was some variation in the amount of persistence that is seen on different visits with the same stimulus exposure. As noted in an early study of persistence (Long et al 2010) and discussed further in a companion report (Long & Baggett 2018), there are some unambiguous examples of intrinsic variations in persistence i.e., cases where both the stimulus exposure and the timing of darks following the stimulus were identical yet the resulting persistence is significantly different. Investigation of these anomalous visits have failed to identify the underlying root cause of the disparity.

One way to handle the possibility that persistence varies intrinsically is to generate calibration files using only a single visit for each exposure time and then to check how well an N x M model removes persistence in those visits. To try this approach, we randomly selected a single visit for each exposure time. We constructed calibration files from these visits, and reconstructed a correction flat. We then ran the same tests as before to assess how well persistence was subtracted. The results are shown in Fig. 8 for a 4 x4 model and summarized in Table 5.

The results from the single visit-based 4x4 model are somewhat better than those obtained from a calibration file generated from all of the visits in Table 1. At least for this particular random dataset, the standard deviation was 0.026 and 0.006 for early and late delay times, respectively, somewhat better than the values obtained for the full set (0.051 and 0.007).

6.2 Measurement Error

A straightforward way to estimate the measurement error is to evaluate the errors at low fluence levels (where one does not expect much persistence) and with a fluence range such that one is sampling about the same number of pixels per cell (500 to 700) as in the high fluence case. Choosing a stimulus range of 15000 to 20000 electrons yielded the requisite number of pixels per cell. The resulting residuals are presented in Fig. 9.

The mean fluxes at each stimulus exposure time are within 0.001 e s^{-1} of zero, and the standard deviations average 0.004 and 0.0035 at early and late delay times, respectively. These numbers are smaller than those for the higher stimulus levels by at least a factor of two, implying that measurement errors are not dominating the measurements. As Fig. 9 shows, there is clearly a cell or two in one of the various exposures that has a larger error, presumably because the fitting procedure with only one data set failed for this particular fluence level.[3]

[3]One of the many trade off studies that could be done to pursue models of this type further would be to optimize the number of levels at which different A, γ models are created. With fewer levels, there would be more pixels to fit within a level.

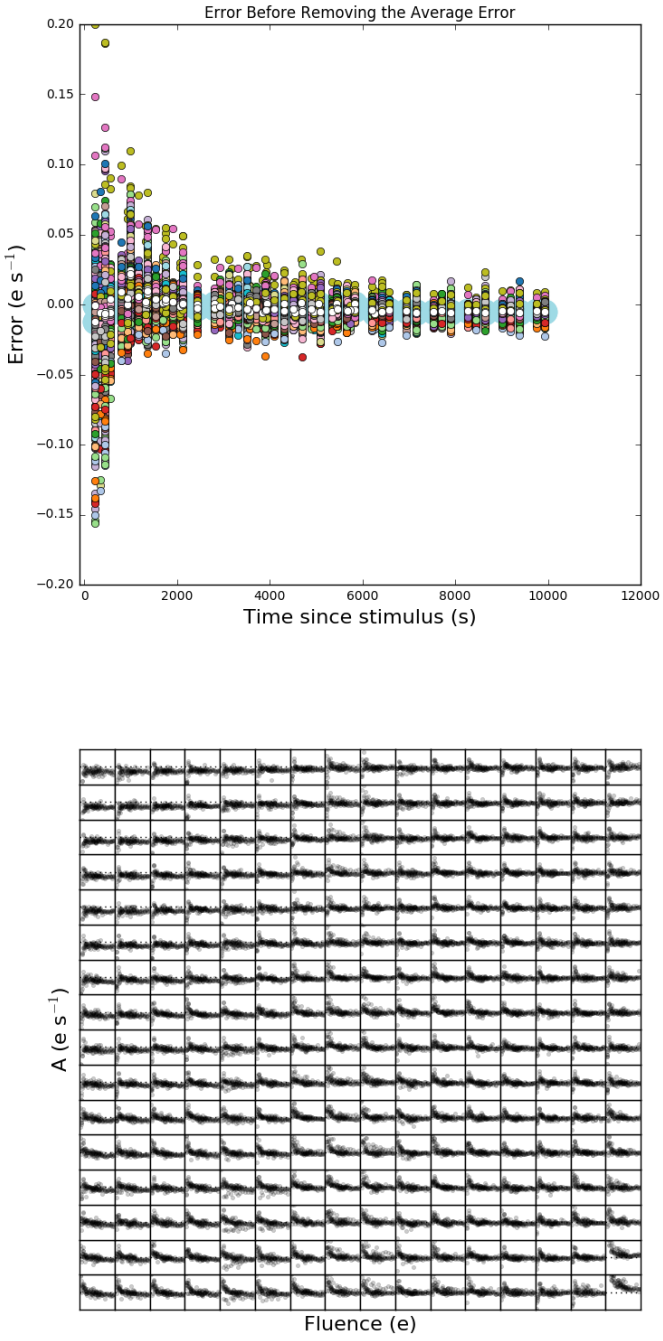

Figure 8. Identical to Fig. 5 and 6, except using only one dataset at each stimulus exposure time.

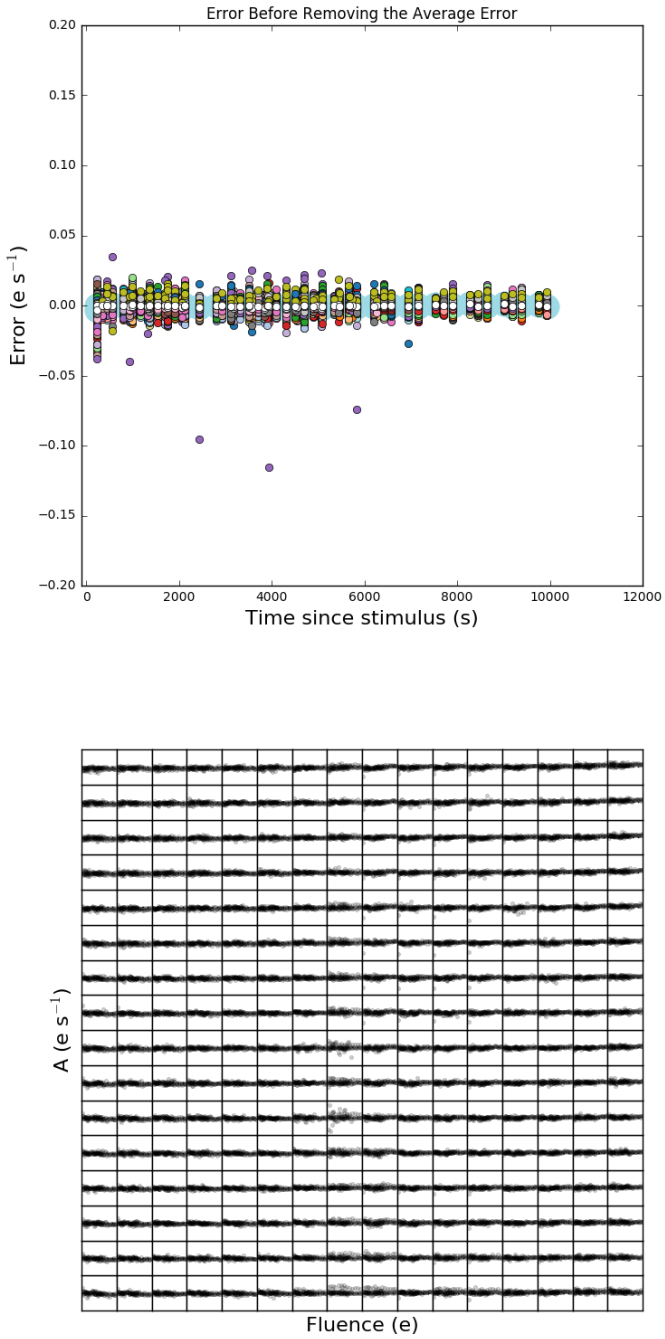

Figure 9. Identical to Fig. 5 and 6, except using only one dataset at each stimulus exposure time. Here we are showing the error in the persistence subtraction when the stimulus levels are low, between 15000 and 20000 electrons.

Table 5. Errors - 4 x 4 Model (One Input Dataset)

Stimulus Exposure	Mean Error (0-1500s)	σ (0-1500s)	Ratio (0-1500s)	Mean Error (>1500s)	σ (>1500s)	Ratio (>1500s)
149	0.006	0.033	0.068	-0.004	0.009	0.128
274	-0.003	0.025	0.022	0.000	0.005	0.044
499	0.007	0.022	0.049	-0.004	0.005	0.082
599	0.008	0.023	0.051	-0.004	0.005	0.082
899	-0.001	0.025	0.020	-0.003	0.004	0.046
1199	0.003	0.027	0.030	-0.003	0.006	0.067

6.3 Departures from a simple power law

For the purposes of this report, we have assumed that persistence in the IR detector decays as a power law for all time. Obviously departures from a power law will contribute to the error. To see whether this is likely to be a factor, we have constructed power law models where the data used for the model is obtained either within 1500 s of the stimulus exposure or more than 1500 s from the stimulus exposure. An example of such fits are shown in Fig. 10. The fit was made using data for the entire array (a simple 1 x 1 model) to reduce the statistical errors as much as possible. The figure is fairly typical of figures generated from data with other stimulus exposure times. The model amplitudes (at 1000 s) are fairly similar, slightly larger when the fit is made to data obtained after 1500 s. The main difference is in the slopes. The power law slopes are generally steeper when one fits the data obtained a longer time after the stimulus image, much steeper at the lower fluence levels, but steeper at long times as well. A steeper slope means that persistence is decaying away more rapidly than one would estimate from the fit at short times.

Errors calculated in the same way as previously are shown in Tables 6 and 7 for the fits using data with delays less than and greater than 1500 s, respectively. Comparing the two tables we see, as expected, that when we fit the data at short delay times we get a better fit to the data at short times than when we use the fit from times with delays greater than 1500 s. Similarly, when we use data at times greater than 1500 s for the fit, we get a better fit for that interval. Furthermore, the fits are better than the corresponding parts of Table 4, which uses a single power law fit to the entire dataset for each visit.

All of this suggests that one could obtain slightly better fits to the persistence if one adopted a more complex model than a simple broken power law.

7 Summary and Conclusions

The IR detector on WFC3, like most other HgCdTe detectors, exhibits persistence when exposed to bright sources, a phenomenon that is associated with defects (traps) in the

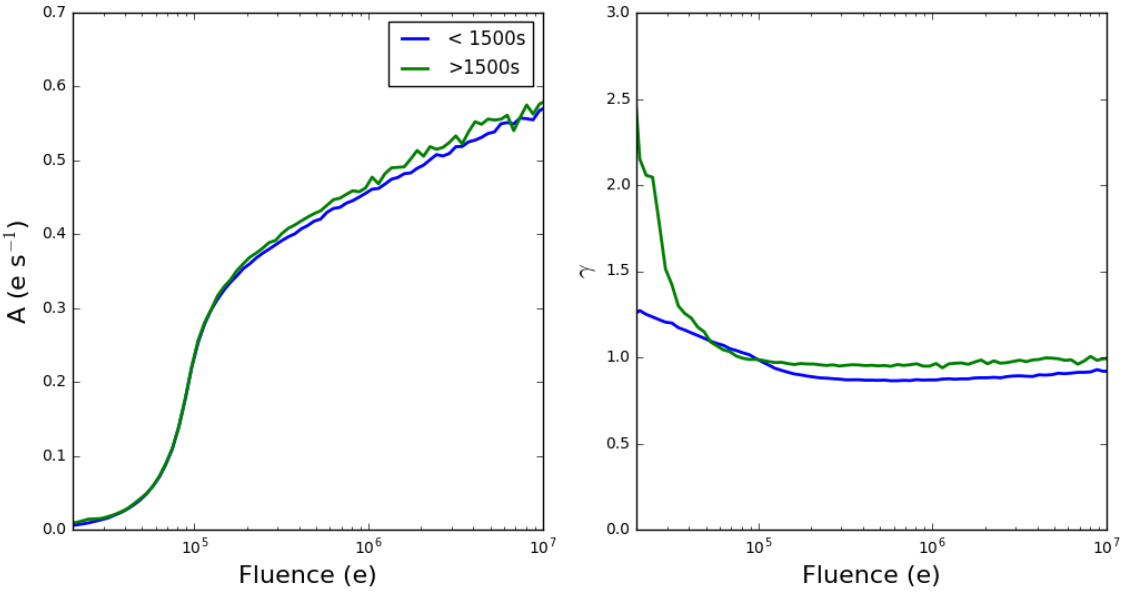

Figure 10. A comparison between power law fits to the data taken within 1500 s of the stimulus image in visit 06 of program 14015 (blue) to a fit to the persistence in darks taken at times greater than 1500 s after the stimulus exposure (green).

Table 6. Errors - Using Data With Less than 1500 s Delay

Stimulus Exposure	Mean Error (0-1500s)	σ (0-1500s)	Ratio (0-1500s)	Mean Error (>1500s)	σ (>1500s)	Ratio (>1500s)
149	-0.008	0.051	0.082	-0.014	0.012	0.210
274	0.006	0.031	0.023	0.002	0.004	0.044
499	-0.003	0.033	0.041	-0.015	0.006	0.193
599	-0.000	0.033	0.041	-0.015	0.006	0.196
899	-0.002	0.053	0.034	-0.006	0.005	0.067
1199	-0.004	0.052	0.046	-0.008	0.007	0.108

Table 7. Errors - Using Data With Greater than 1500 s Delay

Stimulus Exposure	Mean Error (0-1500s)	σ (0-1500s)	Ratio (0-1500s)	Mean Error (>1500s)	σ (>1500s)	Ratio (>1500s)
149	-0.070	0.109	0.124	-0.001	0.011	0.139
274	0.019	0.039	0.029	0.000	0.004	0.039
499	-0.057	0.089	0.080	-0.001	0.005	0.050
599	-0.055	0.089	0.077	-0.000	0.004	0.045
899	-0.058	0.098	0.054	-0.001	0.004	0.032
1199	-0.057	0.087	0.068	-0.001	0.007	0.056

diodes that make up the detector. As we have detailed previously (e.g. Long et al, 2013a, 2013b, 2015a, 2015b), persistence decays as a power law, unlike the exponential decay that is seen in some other IR detectors and unlike the traps one traditionally studies in an elementary quantum mechanics course. The amount of persistence varies slowly across the face of the detector, with higher persistence in the upper left quadrant than in the lower right quadrant.

Researchers using data obtained with the IR channel of WFC3 need to be aware of persistence, as such features can be confused for stars or extended sources in the current image. In order to estimate the amount of persistence, we initially developed a model for persistence based on the assumption that although the power law exponent associated with persistence varied with both fluence and exposure time in the stimulus image, the only variation in persistence characteristics was that the amplitude changed with position on the detector. In this ISR, we have explored a model that relaxes this assumption. Specifically, we have allowed different regions of the detector to have different power law decays. In principle, given enough calibration data one could construct such a model for each pixel of the detector, and thereby characterize persistence for every single pixel (without the need to average over pixels).

However, with only a finite amount of calibration data and read noise limiting the accuracy of measurements made on individual pixels, the model we have explored is one of N X M independently-treated regions across the detector.

The model currently used to predict persistence and produce persistence products for all observers is effectively a 1 x 1 model, though the calibration data used to create this model is not the same as used for the analysis described in this report. The intent was to investigate whether a 2 x 2 or 4 x 4 model, with spatial variations in both amplitude and power law, would improve the quality of the persistence estimation. We found that improvements with such models were modest, if they existed at all and thus, there is no significant benefit at this time to updating the pipeline persistence processing. The primary limiting factors in improving the current persistence model further are the ‘intrinsic” variations between otherwise identical visits and the fluctuations in dark current.

Acknowledgements

We thank George Chapman, Merle Reinhart, Alan Welty, and Bill Januszewski for their help in crafting and executing these programs and Susana Deustua for a careful reading of this manuscript.

References

Hilbert, B & Petro L., 2012, “WFC3/IR Dark Current Stability,” WFC3 ISR 2012-11

Long, K. S., & Baggett S. M., 2018, “Persistence in the WFC3 IR Detector: Intrinsic Variability,” WFC3 ISR 2018-03

Long, K. S., Baggett S. M., Deustua, S., & Riess, A., 2010, "WFC3/IR Persistence as Measured in Cycle 17 using Tungsten Lamp Exposures," WFC3 ISR 2010-17

Long, K. S., Baggett S. M., & MacKenty, J. W. , 2013a, "Characterizing Persistence in the WFC3 IR Channel: Finite Trapping Times," WFC3 ISR 2013-06

Long, K. S., Baggett S. M., & MacKenty, J. W. , 2013b, "Characterizing Persistence in the WFC3 IR Channel: Observations of Omega Cen", WFC3 ISR 2013-07

Long, K. S., Baggett S. M., & MacKenty, J. W. , 2015a, "Persistence in the WFC3 IR Detector: An Improved Model of Persistence for the IR detector on WFC3", WFC3 ISR 2015-15

Long, K. S., Baggett S. M., & MacKenty, J. W. , 2015b, "Persistence in the WFC3 IR Detector: Spatial Variations", WFC3 ISR 2015-16

Smith, R.M., Zavodny, M., Rahmer, G. & Bonati, M., 2008a, "A theory for image persistence in HgCdTe photodiodes," Proceedings of the SPIE, 7021, 70210J-1

Smith, R. M., Zavodny, M., Rahmer, G., Bonati, M., 2008b, "Calibration of image persistence in HgCdTe photodiodes," Proceedings of the SPIE, 7021, 70210K-1

Sunnquist, B., Baggett S., & Long, K. S. 2017, "An Exploration of WFC3/IR Dark Current Variation", WFC3 ISR 2017-04