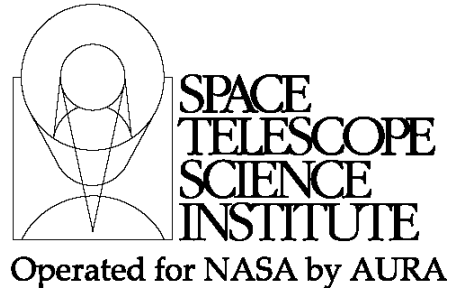




TECHNICAL REPORT



Title: Simple Examples of Simulated NIRCam FITS Data Products	Doc #: JWST-STScI-001650, SM-12 Date: January 16, 2009 Rev: -
Authors: Russell Makidon, Jay Anderson, Robert Jedrzejewski & Roeland van der Marel Phone: 410-338-4946	Release Date: 22 July 2009

1.0 Abstract

We have produced an example FITS data product from a simulated NIRCam MULTIACCUM read sequence on one of the short-wavelength SCAs. Our simulations made use of simplified models of the NIRCam detectors, cosmic ray impacts, photon statistics and detector read noise to produce a series of ten “up-the-ramp” samples of a simulated source. We also created a simulated pipeline to identify cosmic ray impacts to pixels between reads, and to produce a count rate image that simulates the final science data product. We used the example of HST’s NICMOS FITS data products to produce an example NIRCam FITS data product containing one science data extension and four data quality extensions. These products will be available to our JWST Project partners. The goal here was not to make particularly realistic simulations of JWST data, but rather to allow the JWST Project partners to familiarize themselves with the types of data products they may expect from the JWST Data Management Subsystem (DMS) at the Science and Operations Center (S&OC).

2.0 Introduction

The partners of the S&OC on the JWST Project will need to understand the proposed structure of the FITS files delivered to scientists and engineers during JWST Commissioning and Normal Operations. Though the final decisions on the structure of the calibrated data products produced by the Data Management Subsystem (DMS) have not yet been made, one can make educated guesses as to the format of these data files based on our experience with HST data. For the present report, we simulated FITS data products with multiple data extensions, following the example provided by HST’s NICMOS data products.

Though providing our partners with a reasonable FITS file was the goal, producing that file and its associated data required simulating an image (for which we chose an out-of-focus point spread function, or PSF, as expected during wavefront sensing maintenance

Operated by the Association of Universities for Research in Astronomy, Inc., for the National Aeronautics and Space Administration under Contract NAS5-03127

observations with NIRCam’s weak lenses), simulating an infrared detector array read sequence with associated defects, and simulating a simple calibration pipeline. Here we describe each process in turn, and discuss the prospect of extending these simulations in the future to include high-fidelity optical and detector models, along with a set of simulated calibration pipeline processes corresponding to Level 1, 2, and 3 processing of science data (see Section 5.7.3 of the JWST Mission Operations Concept Document, JWST-OPS-002018 Rev B¹).

3.0 Simulations

3.1 Image simulations

We first used the JWPSF² point-source simulation tool (Cox and Hodge 2006) to create a series of point-spread functions (PSFs) that we used as the basis for our simulated data. JWPSF takes as input optical path difference (OPD) maps at the JWST exit pupil and calculates the PSF that results from that OPD map in the absence of photon or detector noise. For this case, we used as our starting point the OPD maps for the Optical Telescope Element (OTE) and NIRCam that met the JWST Revision T (“Rev T”) Optical Error Budget specification, as provided to GSFC via Ball Aerospace. We then modified these OPDs with a parabolic OPD function corresponding to the defocus values expected from the NIRCam weak lenses³ at $\lambda = 2.12 \mu\text{m}$, the primary wavelength of interest for Wavefront Sensing and Control (WFS&C) observations. We then selected a single Rev T OPD and a parabolic OPD function corresponding to roughly +4 waves of defocus, and simulated PSFs at the nominal NIRCam short-wavelength channel pixel scale (0.032” per pixel) from that combined OPD.

NIRCam is comprised of two imaging modules (Modules A and B), both of which make use of short-wavelength ($\lambda = 0.6 \mu\text{m}$ to $2.5 \mu\text{m}$) and long-wavelength ($\lambda = 2.5 \mu\text{m}$ to $5.0 \mu\text{m}$) detectors. Four Sensor Chip Assemblies (SCAs), each of 2048 x 2048 pixels in size, make up one short-wavelength channel focal plane assembly, while the long-wavelength focal plane array consists of a single 2048 x 2048 array. For this work, we simulated a PSF at the center of one short-wavelength SCA. We did not simulate the 4-pixel band of reference pixels lining the edge of each SCA (the total number of light-sensitive pixels in each physical NIRCam SCA is 2040 x 2040).

The PSFs produced by JWPSF are essentially “PSF probability distributions” with a total flux over the entire array equal to 1.0. Thus to simulate a MULTIACCUM sequence of reads of the PSF on a detector, we first calculated a reasonable photon flux for the simulation (approximately $4.0\text{e}+07$ photons/sec), and then multiplied the PSF array by 10 times this flux to simulate a ten-second integration (the shortest integration time available for a full-frame NIRCam SW image). We created ten separate ten-second integrations in this manner, multiplying these integrated PSFs with a simulated detector flat field, and

¹ See also Section 3.1 of Whitman and Valenti (2008)

² JWPSF is available for download at <http://www.stsci.edu/jwst/software/jwpsf>

³ The three NIRCam Weak Lenses defocus the PSF by ± 8 , ± 4 , and $+12$ waves at $\lambda = 2.12 \mu\text{m}$ when used in combination with one another.

adding Poisson noise and cosmic rays (CRs) to each individual array following the methods presented by Anderson (2009). We then summed these integrations to create a MULTIACCUM sequence of detector Reads, e.g. Read 1 = Integration 1, Read 2 = Read 1 + Integration 2, and so forth. In this manner, CRs introduced in previous Reads are preserved, while new CRs are added to subsequent reads. Finally, we added a Gaussian noise model with 15 electrons RMS to each Read to simulate detector read noise.

The CR model we used was a simplified model, both in terms of the frequency and the appearance of the CRs. We assumed a CR impact rate of one per pixel every $\sim 10,000$ seconds. A ten-second integration time between reads suggests any given pixel is impacted by a CR of order every 1000 reads, or that ~ 4200 pixels are impacted by CRs during each integration. Our CRs themselves were divided into two groups: those that hit with an energy equivalent to 500 counts per pixel, and those that hit with an energy of 5000 counts per pixel. It was assumed that each CR affects only a single pixel. In future simulations we plan to include a more accurate CR model that provides both a distribution of CR energies and incidence angles.

Finally, we simulated detector defects in $\sim 1\%$ of the pixels, in accordance with expectations for the HAWAII-2RG IR detectors that will be used for the NIRCcam short-wavelength channels. These defects were represented by “plus signs” in the simulated images (each defect impacted 5 adjacent pixels). Note that with each read, we observe more signal from the source and more cosmic rays. The image defects stay the same. The MULTIACCUM read-sequence data thus obtained is shown in Figure 3-1. In the language of the MOCD, this would be referred to as Level 1 calibrated data.

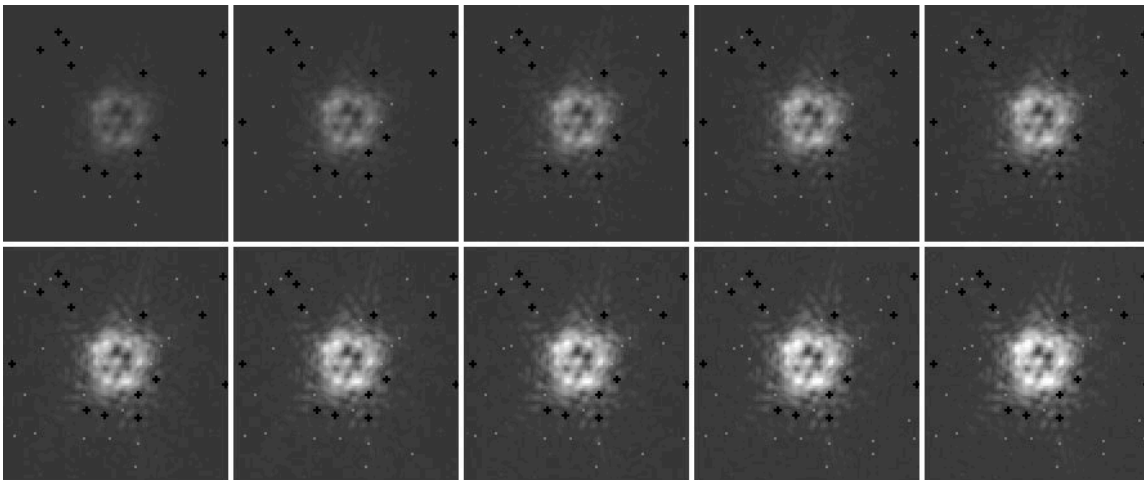


Figure 3-1: A linear MULTIACCUM read sequence for the +4 waves defocused PSF. The PSFs are shown on the same grayscale, with a minimum and maximum display value of 0 and 35,000 respectively. The nominal time between reads is 10 sec. Image defects, which contain 0 counts, are shown as the black + symbols.

Check with the JWST SOCCER Database at: <http://soccer.stsci.edu/DmsProdAgile/PLMServlet>
To verify that this is the current version.

3.2 Image Processing

We developed a simplified set of processes that takes simulated MULTIACCUM reads and produces a science data product similar to what we expect will be produced by the Data Management Subsystem (DMS). While the requirements on the calibration procedures and the structure of the FITS files produced by DMS have not yet been finalized, we have made some assumptions based on prior experience in calibrating HST's NICMOS infrared arrays and the resultant NICMOS data products. A more detailed description of the common calibration processes that we expect to perform on JWST data is available in Robberto et al. (2008).

Our simulated data calibration pipeline followed a different procedure than what is described in Robberto et al. (2008) and in McCullough et al. (2004), where the calibration process requires corrections based on bad pixel masks, reference pixels, bias and dark corrections, and linearity corrections prior to creating a count-rate image (which serves as the science data product in this example). In the actual calibration pipeline, we expect the data will be analyzed pixel-by-pixel by computing a linear fit of the flux in each pixel up to the saturation threshold. Outliers from the fit are identified as CR hits, with a rejection threshold set by the user. When a cosmic ray is found, the reads before and after the cosmic ray event are independently fit by two straight lines and their slopes averaged. This procedure can be iterated to look for multiple events, with the number of iterations set by a dedicated keyword.

In the present application, we fit a slope to each pixel from the individual up-the-ramp frame reads using a "quick-and-dirty" method. Here we took the differences between the pixel values for successive reads and calculated a sigma-clipped mean of these differences. This allowed for a cursory rejection of CRs, as shown in Figure 3-2, and enabled us to calculate a reasonable count-rate image as our science data product. We also used the information obtained from this process to determine the RMS error in each pixel based on an analysis of the error in the average of the differences in each pixel. This pixel error array is saved and delivered as an extension in the FITS data product, as described in Section 4.0.

The count-rate images thus obtained is shown in Figure 3-2. In the language of the MOCD, this image and the associated data products described in Section 4.0 would jointly be referred to as Level 2 calibrated data.

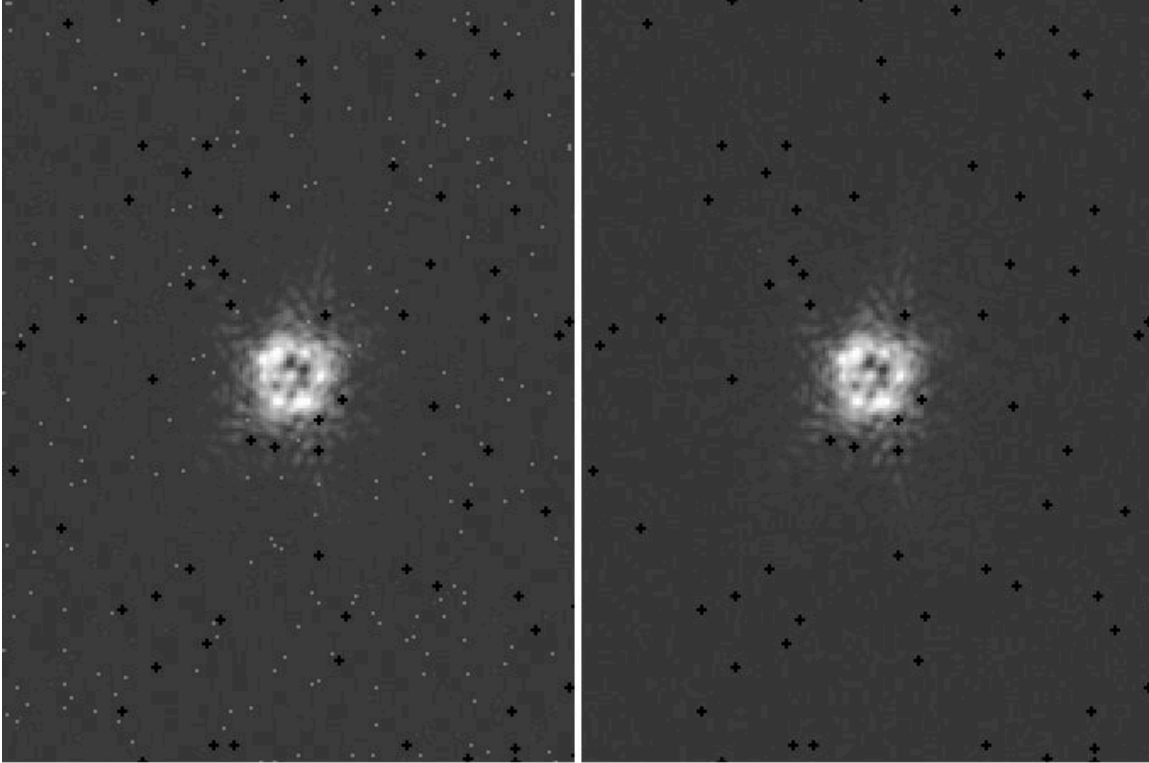


Figure 3-2: Comparison between the 10th read in the simulated MULTIACCUM sequence (left) and the count-rate image calculated from the differences in the MULTIACCUM reads (right). Note that while most of the CRs have been identified and removed in the count-rate image, the image defects remain.

3.3 Artifacts Remaining After Image Processing

In the case described above, our models of CR flux and CR impacts on the image enabled us to use simplified CR detection methods in image processing. While the DMS will undoubtedly use a more robust method to detect CRs with MULTIACCUM exposures (see Robberto 2008, Offenberget al. 2005), identification of CRs in real image data may be complicated by a number of factors. For example, our simulated CRs were easy to detect because, by construction, all of the simulated CRs stand out well above the noise in the image (see Figure 3-2). For actual observations, CRs exhibit a spectrum of energies, and may impact the detector in very different ways (e.g., from low-energy, single-pixel events to high-energy impacts that affect multiple pixels at one time). Because of this, some positive outliers, i.e. poorly or partially corrected cosmic rays, might still remain in the post-processed images. Such artifacts will occur at different positions in each image obtained with the instrument, and would not be known prior to image processing.

Each infrared array used on-board JWST will likely have a small number of bad pixels that react differently to photons than the typical pixels in the array. These pixels can exhibit a very low or zero response (“cold” pixels, as in the case with these simulations), or show very high or erratic dark currents (“hot” pixels). If these bad pixels are known

Check with the JWST SOCCER Database at: <http://soccer.stsci.edu/DmsProdAgile/PLMServlet>
To verify that this is the current version.

and their characteristics are stable, then these pixels can be “flagged” as bad and their positions marked in the Data Quality array (see Section 4.0 below). However, as experience with HST’s CCD and NICMOS arrays have shown, the number of bad pixels in a detector array grows with time, a result of the damage the arrays suffer in a high-radiation environment. Transient effects have also been noted, where pixels act bad for a time and then appear to correct themselves. New or transient defects would not be “flagged” by the DMS during image processing, and as such would not appear in any data quality description.

Even when a defect is “flagged”, the DMS pipeline will never perform an interpolation over the defect to “correct” those pixels. If such interpolations are desired, they will need to be performed by the users in post-calibration processing.

We also note that the prior exposure history of an infrared array can impact subsequent exposures with that array. High flux levels - from a source or from CRs - that cause pixels to collect large amounts of charge will create an elevated dark current in that region of the array, essentially leaving remnant images of the high-flux source for a series of exposures. Though the mechanisms that cause this “persistence” are not well understood (Smith et al. 2008a), it appears that these residual images decay on multiple timescales in Hawaii-IIRG arrays like those to be used in JWST instruments (Smith et al. 2008b). Understanding image persistence remains one of the challenges of infrared astronomy, and as such it is unlikely pixels impacted by image persistence will be flagged by the DMS during routine image calibration.

For the stated reasons, additional processing is usually required to further improve a count-rate image, such as that shown in Figure 3-2, for scientific use. Such processing would typically use dithered data, i.e. multiple datasets of the same scene with a small shift on the detector. Combination of such datasets can further reduce the impact of both known and transient image artifacts. This is described in more detail in Anderson (2009b). In practice, the DMS may use software such as MultiDrizzle (Fruchter, A. and Sosey, M. et al. 2009) to combine dithered data. In the language of the MOCD, calibrated data obtained through the combination of multiple datasets with different pointings would be referred to as Level 3 calibrated data.

4.0 FITS file product

As noted above, the final details of the data products produced by the DMS have yet to be determined. However, we do expect to take advantage of the experience gained from processing and delivering science data products for HST’s NICMOS. The primary science data product for NICMOS is a single FITS file with five arrays, each stored as a separate image extensions in a FITS file⁴. We follow that structure here, noting that in reality the actual NIRCcam FITS files produced by the DMS may contain more than the five data extensions listed below.

⁴ A full description of the NICMOS data pipeline and science data products is available in the NICMOS section of the HST Data Handbook, available at:

<http://www.stsci.edu/hst/nicmos/documents/handbooks/DataHandbookv7/>

Check with the JWST SOCCER Database at: <http://soccer.stsci.edu/DmsProdAgile/PLMServlet>

To verify that this is the current version.

- Extension 1 is the science image array (SCI), count-rate image formed from the MULTIACCUM read sequence.
- Extension 2 is the Error image, which is a floating-point array containing an estimate of the statistical uncertainty associated with each corresponding science image pixel. In the HST NICMOS example, this array is computed by the ground calibration pipeline based on computed uncertainties in the count rates derived from the linear fit to counts v. exposure time from the intermediate readouts in a MULTIACCUM read sequence. Here we compute the error based on an analysis of the error in the average of the differences in each pixel between simulated detector Reads.
- Extension 3 is the data quality (DQ) array. This is an integer array containing bit-encoded data quality flags indicating various status and problem conditions associated with corresponding pixels in the science image. Here we use a value of 0 to denote "No known problems", while a value of 32 signifies a defective (hot or cold) pixel. CR-affected pixels are assigned a value of 512. These values correspond to the same data quality flags used in the NICMOS data products. In reality, some defects will not be marked because they may be new or transient or unknown to the system, and some CRs may not be recognized by the processing pipeline.
- Extension 4 is the number of data samples (SAMP) array, an integer array containing the total number of data samples that were used to compute the corresponding pixel values in the science image. Cosmic rays that saturate the pixels, for example, will result in a reduced number of reads for that pixel.
- Extension 5 is the integration time (TIME) array, containing the effective integration time associated with each corresponding science image pixel value in a floating-point format.

In the real case, each of the five image arrays would have header information specific to that array. For example, the SCI extension would contain information regarding the World Coordinate System, providing the relationship between the detector pixels and on-the-sky coordinates, along with photometry keywords, image statistics, and data quality information. Global information pertaining to all data in the file extensions, such as information specific to the proposal, would be kept in the primary extension (Extension 0, for which there is a header, but not an associated data product). In this case we did not include any extension-specific header information in the data products described here, though we did populate the primary extension header with basic information about our simulations.

5.0 Future Work

The focus of this work was to produce a FITS data product with a structure similar to what we might expect from NIRCcam image data and not to produce a high-fidelity image simulation or calibration pipeline. However, we expect that future work in this area will

Check with the JWST SOCCER Database at: <http://soccer.stsci.edu/DmsProdAgile/PLMServlet>
To verify that this is the current version.

require more realistic representations of both the detector and the calibration processes. Future simulated NIRCAM data products should include data from all four short-wavelength SCAs if those SCAs will be packaged together (as with the Hubble ACS WFC images) as well as fully-populated FITS headers for all available data extensions.

The final simulated data products produced for this report are available, along with a brief README file describing the contents of the FITS files, at the following URL: http://www.stsci.edu/~jayander/NIRCAM_SIM_PROD.tar.gz.

The S&OC at STScI is currently evaluating the state and viability of existing image and detector simulation code with a goal of producing highly accurate science data simulations. In addition, detailed requirements on the DMS and its data products will also start to be produced in the near future. We expect to continually improve our simulated data products as our knowledge of the instruments, detectors, and calibration procedures improve.

6.0 References

- Anderson, J. 2009, “JWST_PSSIM: A Point-Source Simulator for JWST”, (JWST-STScI-001587)
- Anderson, J. 2009, “Use and Benefits of Dithered WFS Observations”, (JWST-STScI-001722).
- Cox, C. and Hodge, P. 2006, “Point-spread function modeling for the James Webb Space Telescope”, Proc. SPIE, Vol. 6265, p. 62650W
- Fruchter, A. and Sosey, M. et al. 2009, “The MultiDrizzle Handbook”, version 3.0 (Baltimore:STScI)
- Long, K. 2006, “JWST Mission Operations Concept Document”, (JWST-OPS-002018)
- McCullough, P., Figer, D., Rhoads, J., Kelly, D. & Rieke, M. 2004, “NIRCAM Science Data Pipeline Description”, (JWST-STScI-000381)
- McLaughlin, H. and Wiklind, T. 2007, “HST Data Handbook for NICMOS”, Version 7.0 (Baltimore: STScI)
- Offenberg, J. D., Fixsen, D. J., and Mather, J. C. 2005, “Memory-Efficient Up-the-Ramp Processing with Cosmic-Ray Rejection”, PASP, 117, 94
- Robberto, M., McLaughlin, H., Beck, T., Fullerton, K., Gordon, K. Jedrzejewski, R., & Regan, M. 2008, “Concept of a common Science Data Pipeline for JWST and Initial CALWebb Calibration”, (JWST-STScI-001428)
- Robberto, M. 2008, “On the Optimal Strategy to Fit MULTIACCUM ramps in the presence of cosmic rays,” (JWST-STScI-001490)
 Check with the JWST SOCCER Database at: <http://soccer.stsci.edu/DmsProdAgile/PLMServlet>
 To verify that this is the current version.

- Smith, R. M., Zavodny, M., Rahmer, G. and Bonati, M. 2008, "A theory for image persistence in HgCdTe photodiodes", Proc. SPIE, Vol. 7021, p. 70210J
- Smith, R. M., Zavodny, M., Rahmer, G. and Bonati, M. 2008, "Calibration of Image Persistence in HgCdTe Photodiodes", Proc. SPIE, Vol. 7021, p. 70210K
- Whitman, R. & Valenti, J. 2008, "Data Management Subsystem Operations Concept Document (DRAFT)", (JWST-STScI-001526)