STScI | SPACE TELESCOPE SCIENCE INSTITUTE

JWST TECHNICAL REPORT

Title: NIRISS Commissioning Results: NIS-011a Illumination Flat Fields	Doc #: JWST-STScI-008297, SM-12 Date: 31 October 2022 Rev: -
Authors: Jo Taylor, Kevin Volk, Chris Willott, Paul Goudfrooij, André Martel, & Tony Sohn Phone: (410) 338-4501	Release Date: 07 November 2022

1 Abstract

We summarize the L-, or low-frequency, flat field effects of the Near-Infrared Imager and Slitless Spectrograph (NIRISS) detector during James Webb Space Telescope (JWST) commissioning. Data from program 01086 (NIS-011) were used to derive new delta L-flats, which were then incorporated into new flat reference files. We find that the average delta flat ratio across the NIRISS detector agrees with the ground flats. For a subset of pixels there is, at maximum, a 7.8% change in on-sky flats compared to ground flats. We also discuss the limitations to fitting the low-frequency component of the flat fields when on-sky pixel-to-pixel flat field changes are still unknown.

2 Data and Calibration

NIRISS program 01086 (Martel et al. 2020) serves three purposes: 1) to measure the detector distortion coefficients, 2) to verify subarray photometry, and 3) to evaluate the large-scale response variations across the detector when it is uniformly illuminated. The flat-field determination aspect of program 01086/NIS-011 is referred to as NIS-011a. The ISIM (Integrated Science Instrument Module) source used in cryo-vacuum 3 testing (CV3) did not necessarily provide uniform illumination across the NIRISS detector, therefore any large-scale variation across the detector must be determined on-sky. There is no single astronomical source that provides uniform illumination over the full field of view, so instead we obtained direct imaging of a dense field of stars dithered across the detector in order to derive relative response values. We chose to observe the 5′x5′ Hubble Space Telescope (HST) astrometric field in the LMC (Large Magellanic Cloud) located in the CVZ-South (Continuous Viewing Zone-South). This L-flat determination method has been employed in the past by the HST/WFC3 and HST/ACS teams (Mack et al. 2002, Mack et al. 2013) and relies on a matrix-solution Fortran program written by Roeland van der Marel (van der Marel 2003). The L-flat correction can be expressed as an over-determined matrix equation with a minimum chi-square solution (Mack et al, 2013).

All twelve NIRISS filters were used in program 01086: F090W, F115W, F140M, F150W, F158M, F200W, F277W, F356W, F380M, F430M, F444W, and F480M. For each filter, we obtained a 3x3 position mosaic with 75% overlap in rows and columns (i.e. each mosaic X and

Y offset corresponded to 512 pixels or ~33”). The mosaic covered a field approximately 3.3′x3.3′ (see Figure 1). At each mosaic pointing, two images were taken using the WFSS MEDIUM 2-point dither pattern, in order to help correct for bad pixels on the detector (each dither X and Y offset corresponds to 6.5 pixels, or 0.425” and 0.426” respectively). In total, 18 individual exposures were obtained for each filter. The ramps had 5 groups for the short wavelength filters (F090W, F115W, F140M, F150W, F158M, F200W), 10 groups for the long wavelength wide filters (F277W, F355W, F444W), and 15 groups for the long wavelength medium filters (F380M, F430M, F480M). For all filters, the number of integrations was 1 and the readout mode was NISRAPID. We refer to the central pointing (i.e. XOFFSET=YOFFSET=0) as the “reference image” for each filter. A summary of the exposure information and star measurements is shown in Table 1.

Table 1: The total exposure time, number of images, average number of stars, and average FWHM for each NIRISS filter. *The average number of stars is after imposing restrictions on star count rate, FWHM, roundness, and separation.

Filter	Total Exposure Time [s]	Total N_{images}	Avg. N_{stars} per image*	Avg. FWHM [pixels]
F090W	966	18	8240	1.364
F115W	966	18	7429	1.377
F140M	966	18	5832	1.420
F150W	966	18	7221	1.434
F158M	966	18	5557	1.443
F200W	966	18	6173	1.477
F277W	1932	18	3884	1.501
F356W	1932	18	2706	1.615
F380M	2898	18	925	1.680
F430M	2989	18	792	1.775
F444W	1932	18	1802	1.781
F480M	2989	18	720	1.853

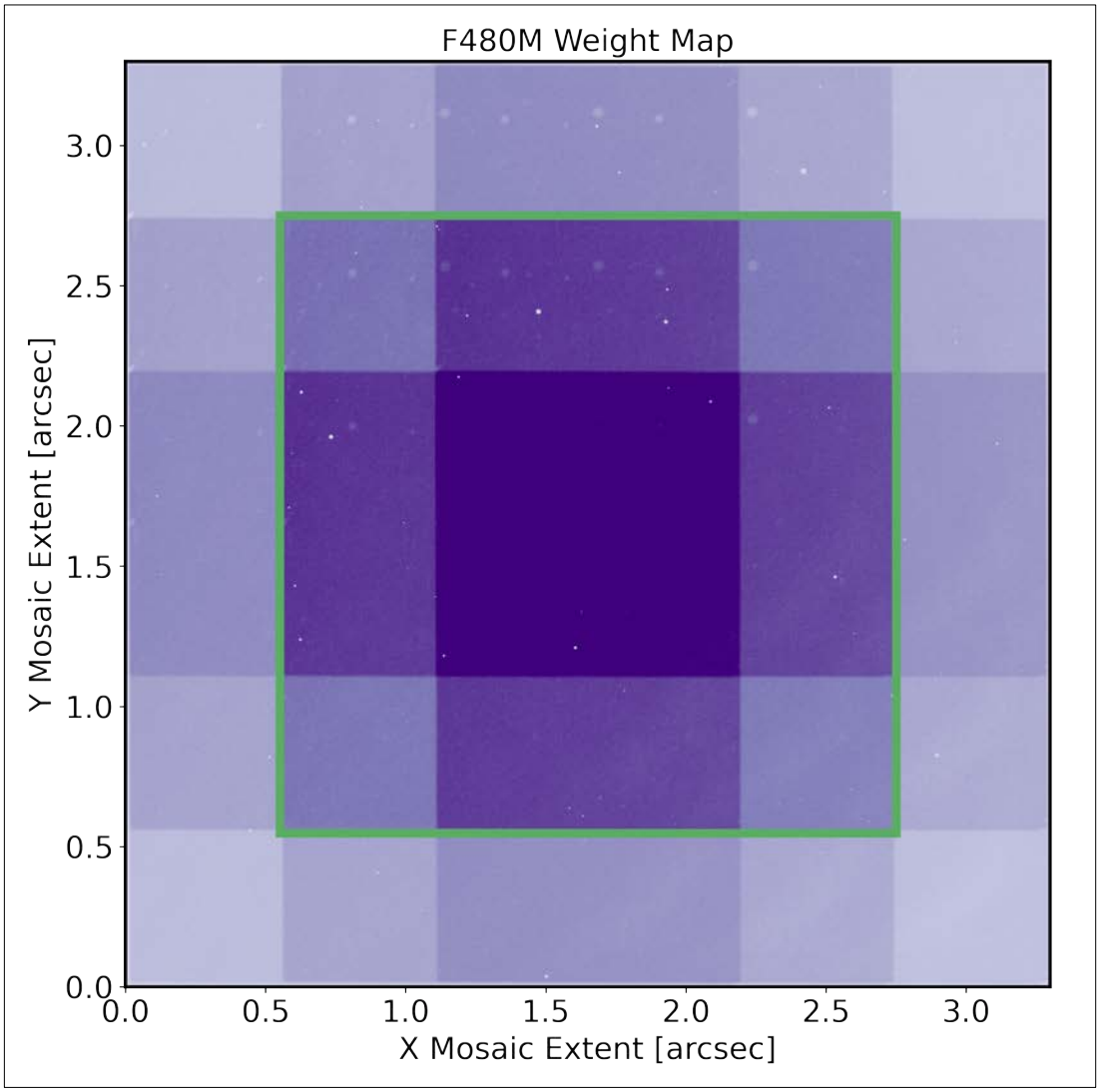

Figure 1: Weight map of F480M drizzled _i2d product. The full mosaic is ~3.3x3.3′, and the extent of the NIRISS detector (2.2x2.2′) is shown in green. Darker colors correspond to regions with more contributing exposures, while lighter colors correspond to regions with fewer contributing exposures.

Multiple rounds of calibration and analysis were performed during commissioning, but from this point on we only describe the methodology used to obtain the final products. The goal of this program was to model the delta L-flat. A delta flat refers to any additional flat field variation that is present *after* calibrating data with a flat reference file. An L-flat models only the low-frequency flat field effects, as opposed to a P-flat which models the pixel-to-pixel variation.

01086 data were calibrated using the JWST pipeline version 1.5.2, with the following modifications.

- In Detector1Pipeline: a preliminary bad pixel mask was used (which was subsequently delivered to CRDS as jwst_niriss_mask_0016.fits).
- In Image2Pipeline: the resample step was skipped, and modified CV3 flat reference files were used. These files were created by Chris Willott to account for evolution of the flat field structure due to time and temperature changes. First, high S/N internal flats

from program 01084 (NIS-009, Cooper et al. 2022) were compared to similar CV3 flats and the ratio applied to the CV3 flats. Then, for imaging flats only, evolution in the pick-off mirror features were updated using data from program 01063 (NRC-010, in prep).

- In Image3Pipeline: in the tweakreg step, the align_to_gaia option was enabled and for each filter, the average FWHM measurements was used for the kernel_fwhm parameter (see Table 1). The average FWHM measurements for each filter were likewise used in the kernel_fwhm parameter in the source_catalog step. In the skymatch step, skymatch.lower- the lower limit of usable pixels- was set to zero. In the resample step, a Gaussian kernel was used rather than a square kernel. Finally, intermediate _crf, or cosmic ray flagged, files from the outlier_detection step were saved for use in further analysis as explained in the following section.
- CRDS context jwst_0863.pmap was used for all reference files except the aforementioned manually supplied mask and flat files.

3 L-Flat Field Measurement

Measuring the L-flat comprises several steps:

1. Calibrating data to obtain Stage 3 products (described in the previous section).
2. Identifying and filtering sources.
3. Performing photometry on each source.
4. Cross-matching sources from each mosaic and dither position.
5. Additional filtering of sources.
6. Fitting a smooth function to the low-frequency flux deviations.

All code used to derive the L-flats is stored on the spacetelescope niriss-commissioning repository.

Source identification and filtering

Sources were identified by Tony Sohn using the centroid_2dg function in photutils (Bradley et al. 2022), which calculates centroids by fitting 2D Gaussians to the data. This method is time-consuming, but proved the most accurate for NIRISS where the majority of filters are undersampled (JWST User Documentation, 2016-2022). Accurate centroids are critical to the delta flat analysis, since cross-matching of sources is performed by comparing sky coordinates and the observed field is quite crowded. Source identification was performed on the _crf files, so we chose to also perform our flat field measurements on the _crf files rather than the level2b _cal products.

After identification, sources were filtered based on multiple criteria: count rate, FWHM, roundness, and source separation. Stars with saturated cores were also excluded from analysis. After filtering, the number of sources per filter is summarized in Table 1. The sources used for analysis are shown for filters F090W and F480M in Figure 2.

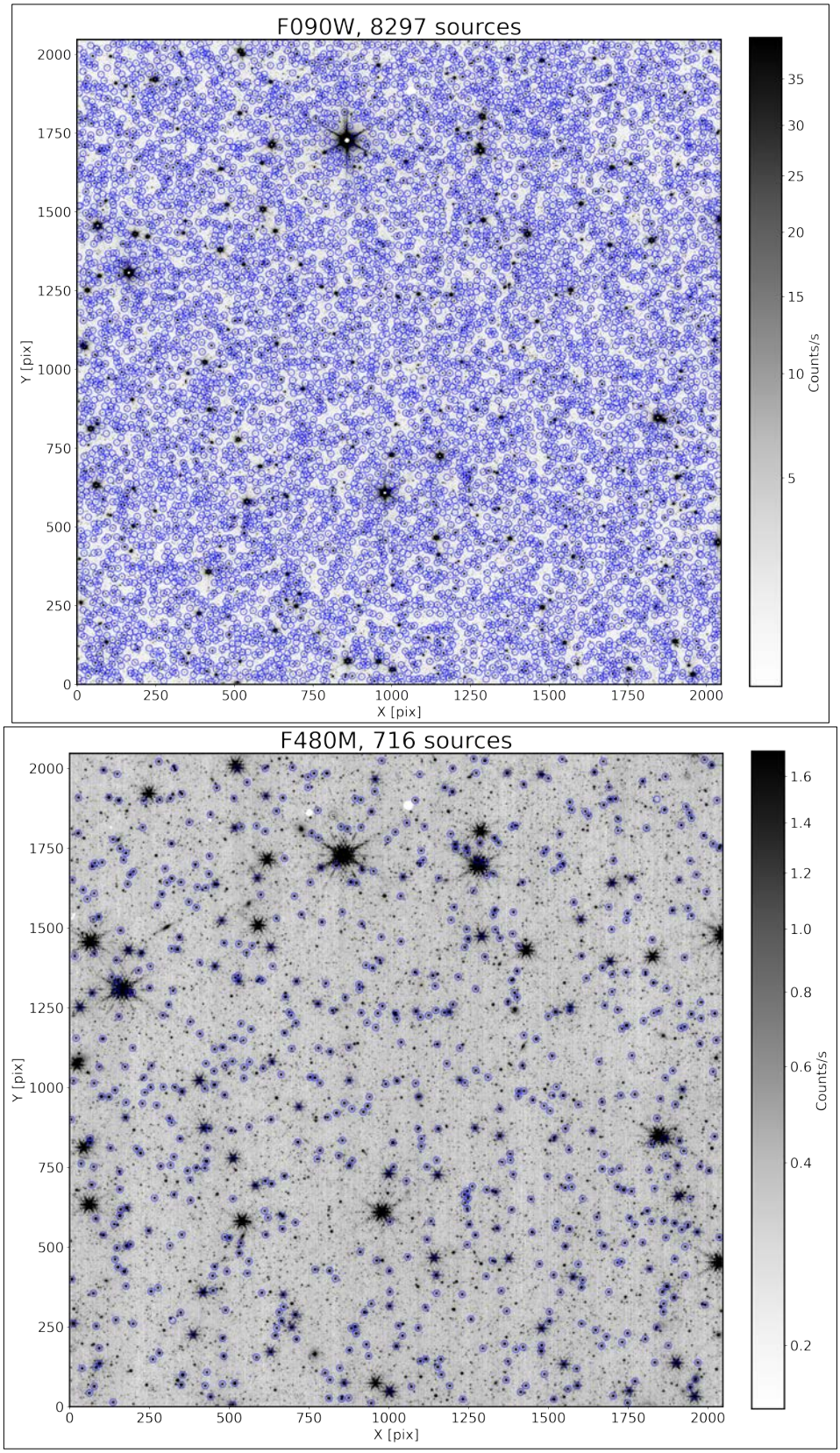

Figure 2: The positions of sources after imposing restrictions on star count rate, FWHM, roundness, and source separation. The reference (central pointing) images for filters F090W (top) and F480M (bottom) are shown.

Table 2: For each filter: the average number of stars per image *after* restrictions on star count rate, FWHM, roundness, and source separation; the average number of stars per image used to perform photometry; the magnitude limit used to select stars for L-flat fitting; and the *total* number of stars used for L-flat fitting.

Filter	Avg. N_{stars} Per Image	Avg. N_{stars} for Photometry	Magnitude Limit [Sirius mags]	N_{stars} for L-flat fitting
F090W	9391	8240	20.31	6413
F115W	8376	7429	20.293	6404
F140M	6243	5832	19.344	7909
F150W	7966	7221	19.83	6224
F158M	5982	5557	18.811	5438
F200W	6893	6173	19.188	4155
F277W	4314	3884	19.293	4256
F356W	2900	2706	19.196	3746
F380M	941	925	16.695	1838
F430M	812	792	15.985	1564
F444W	1832	1802	18.51	2592
F480M	731	720	16.71	1442

Photometry

With a list of preliminary sources in hand, we then performed circular aperture photometry on all stars by using photutils. For the long-wavelength filters in the filter wheel, it was possible to approximate the best aperture size by using a variety of sizes and comparing the resulting S/N ratios. For the shorter-wavelength filters in the pupil wheel, however, field crowding made this impossible. We used an aperture radius of 3 pixels for the short-wavelength filters (F200W and shorter) and 5 pixels for the long-wavelength filters (F277W and longer). For all filters, we used a circular sky annulus with an inner radius of 15 pixels and an outer radius of 20 pixels. We used a custom python script from Paul Goudfrooij to not only perform, but also improve the photometry. This script interpolates over bad pixels within the science aperture, which can greatly improve the results when working with individual exposures and not dithered combined images. Finally, star fluxes were converted from count rates to instrumental magnitudes.

Source matching and final filtering

To perform star matching, we used a python script created by Kevin Volk. This returns a catalog for each filter listing every star's magnitude and error as a function of detector position and exposure number. This catalog is then further reduced by selecting only the best stars for final L-flat fitting. Two criteria are evaluated automatically by the Fortran routine: each star must have 4 or more measurements (due to the 2-point dither pattern, each star is guaranteed to have at least 2 measurements), and each star's RMS error must be less than 0.05 or 3 times the photometric error (whichever is larger). We imposed an additional requirement that each star's average magnitude (and therefore error) must be minimized as much as possible given the number of sources. For the short-wavelength filters, limiting the sample by average error left us with a very large sample of stars for final flat-field fitting. However, there are fewer bright, clean sources in the long-wavelength filters, so relaxed error limits were necessary to obtain catalogs with enough stars to sufficiently cover the full detector space. It is imperative that the final list of stars/measurements span the full detector as much as possible in order to provide robust fitting at the edges and corners. The final number of sources and associated magnitude limits are provided in Table 2. These magnitude limits are supplied to the Fortran routine at runtime.

Flat-field fitting

We now have the best possible catalog of sources for each filter, which we provide to the Fortran code along with several run-time arguments that determine exactly how to fit the flat field. This code was provided to us by Jennifer Mack and only minimal changes were made to accommodate NIRISS data, including increasing the original array sizes. The fitting algorithm is described in detail in van der Marel 2003: the L-flat is expanded as a linear sum of two-dimensional basis functions, and magnitude differences between measurements of the same star at different detector positions constrain the coefficients of the linear sum. This reduces mathematically to an over-determined linear least-squares problem that can be solved through singular-value decomposition (SVD) (van der Marel 2003). As such, the L-flat correction is an over-determined matrix equation with a minimum chi-square solution (Mack et al, 2013). This solution is represented either by a series of Legendre polynomials or as a grid of basis functions. The order of the polynomials/grid of basis functions is specified by the user.

Generally speaking, the polynomial solution is preferred in cases with many stars, many measurements per star, and where star measurements are spread approximately uniformly across the detector. In other scenarios, the grid of basis functions- the chessboard solution- may prove superior. Additionally, the exact nature of the flat variation can influence which solution should be used. The polynomial solution works well for low-frequency variations, but may not be appropriate for all medium-frequency variations, even when using higher-order solutions. In such cases, the chessboard solution may be preferred. Each filter to be fit must be evaluated individually to determine the optimal algorithm parameters. We found that 4th order solutions provided the most accurate results.

We used the following parameters for fitting:

- For filters F380M and F430M only, a very small constant systematic error floor of 0.01 mags was added in quadrature to the catalog errors. The photometric errors in these filters were small enough that robust fitting could not be performed without adding a buffer.
- CTE and zeropoint fitting was disabled.
- Appropriate zeropoints for each filter were provided at run-time.
- The sample of stars to be used for analysis was drawn by imposing a magnitude limit, which varied as a function of filter (see Table 2).
- A 2-D 4th order polynomial basis function was used for all filters.
- Input errors were not multiplied by a fudge factor.
- Residuals were binned into a 16x16 image where each “pixel” was a 128x128 pixel bin.
- The output flat field and error images were sampled to the native NIRISS detector space, i.e. the output images were 2048x2048 pixels.

It should be noted that program runtimes could be quite long, and that they increase greatly as a function of number of stars/measurements. We used the STScI internal witserv machines and the sparsest filter (F480M, ~1400 stars) took 9 minutes, while the densest filter (F140M, ~7000 stars) took 42 hours. Additionally, when compiling the Fortran code on Linux systems, it was imperative to use the -mcmodel=medium argument, otherwise the code could not create a sufficiently large matrix.

The output delta L-flat FITS files from the Fortran code are in units of delta magnitude, which we then convert to ratios. E.g. a delta magnitude of 0 corresponds to a ratio of 1.0 and a delta

magnitude of -0.1 corresponds to a ratio of ~1.097 (level2a data are *divided* by flat field ratios). As a reminder, because we derived the L-flats using data already calibrated with a flat reference file, the resulting fitted flat is actually a *delta* L-flat. We multiply the derived delta L-flats by the flats originally used to calibrate the NIS-011a data (described in Section 2). The product then represents the original flat field with an additional L-flat component folded in. These new flat reference files are then normalized, where the normalization factor is the average of all non-reference pixels that also have good DQ (data quality) values. The error array in the original flat reference file is multiplied by the derived delta L-flat ratios as well. Finally, the new flat reference files are updated to conform to CRDS standards.

4 Results and discussion

The derived delta L-flats and residuals for each filter are shown in Figures 3 and 4, respectively. The change in on-sky flats compared to ground flats is within a few percent for most pixels. For a subset of pixels there is, at maximum, a 7.8% change in on-sky flats compared to ground flats, with the corners and edges of the detector having slightly higher percent differences. This is typical for flat measurements of this type, as there are fewer star measurements in the outer perimeter of the detector due to the mosaic pattern (see Figure 1).

The Fortran code reports the residuals for each measurement of each star. For each measurement, the residual is defined as the measured magnitude minus the predicted magnitude. The predicted magnitude is defined as the "true magnitude" plus a correction accounting for the L-flat at that measurement's position. The "true magnitude" is calculated as part of the matrix SVD.

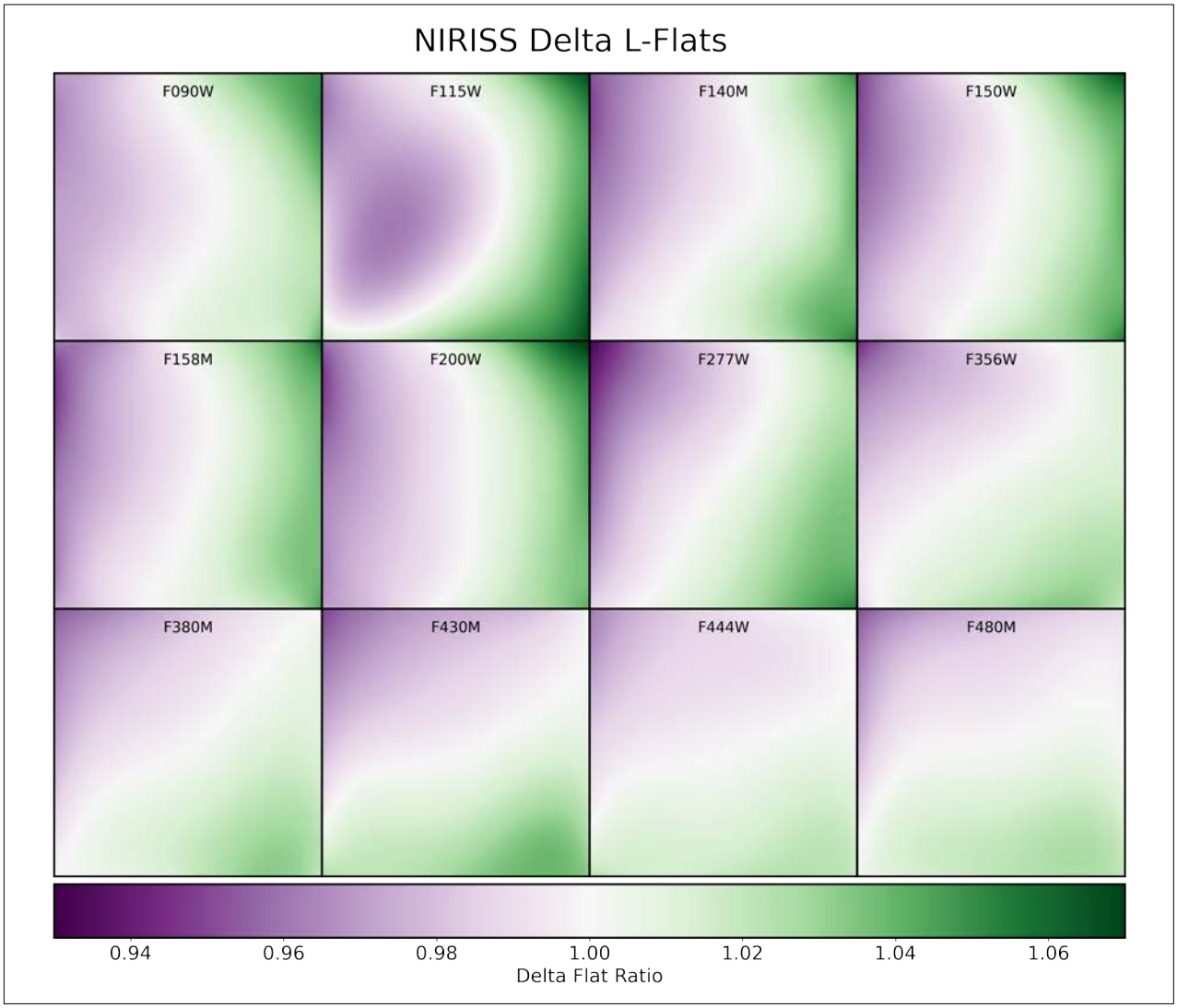

Figure 3: Delta L-flats for all NIRISS filters. Each subplot spans the area of the NIRISS detector. White corresponds to a delta flat ratio of 1.0

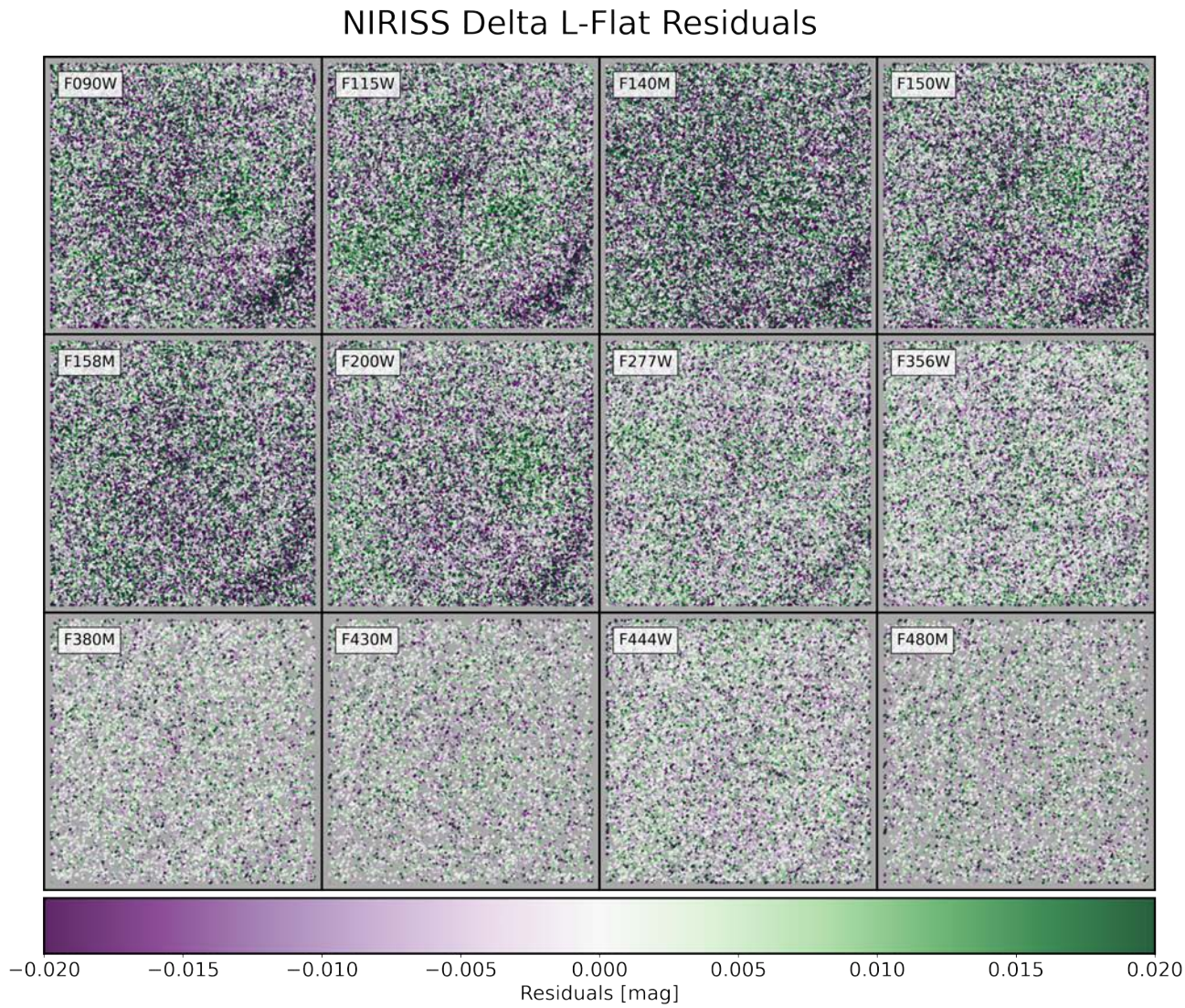

Figure 4: Delta L-flat residuals for all NIRISS filters. Each subplot spans the area of the NIRISS detector. White corresponds to a residual of 0.

There are two methods of evaluating the goodness of fit: examining the residuals, and calibrating data with the new flat reference files then re-deriving the delta L-flat. In the latter case, the so-called "double delta flat" should be ~uniformly 1.0 across the detector, if an adequate original fit was performed. These double delta flats were indeed ~1.0 across the detector.

When examining the residuals however, it is quite obvious that there are systematic trends across the detector (Figures 4 and 5). These spatially dependent residuals are the result of known NIRISS detector features that cannot be modeled at the low-frequency level. The area near the lower right corner of the detector (A, in Figure 5) corresponds to the region of strongest cross-hatching. Cross-hatching effects vary at the subpixel level and will disproportionally affect undersampled PSFs; for NIRISS, filters <4 microns are undersampled (JWST User Documentation, 2016-2022). The more centrally located region (B, in Figure 5) corresponds to the detector "epoxy void". This is an area of the detector that does not contain

any epoxy within the detector. While the void edges do evolve with time, there is no current evidence that the void interior itself has. We do not fully understand why the residuals in the void region are elevated. The void affects the IPC (inter-pixel capacitance), so there may be a potential signal discontinuity at the void edge depending on the size of the photometric aperture used. It is possible this effect could impact the goodness of the fit across the void edges and within the void itself. When attempting to fit a smooth function between these regions A and B, the area in between them is affected as well, resulting in higher than average residuals in region C (Figure 5).

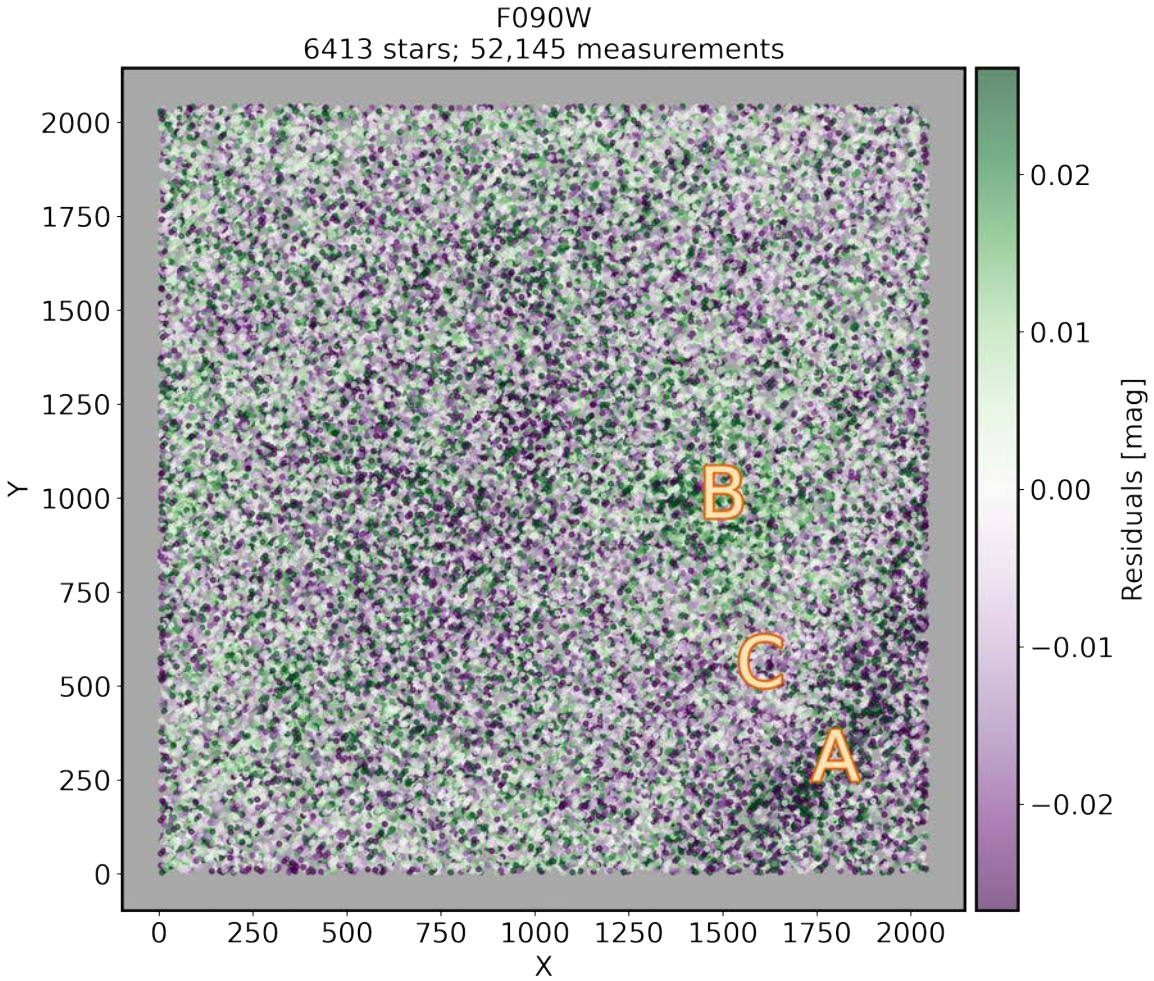

Figure 5: Delta L-flat residuals for the F090W filter. Region A is the area of strongest detector cross-hatching, B is the epoxy void, and C has higher residuals due to polynomial fitting constraints.

As an alternative, we also attempted using a 4th order chessboard solution which essentially computes a pixelized flat field. A 4th order solution corresponds to 2^4x2^4, or 16x16, output flat field, where each superpixel represents a 128x128 pixel box on the native NIRISS detector. This pixelized flat was then up-sampled to the native detector size of 2048x2048 pixels and smoothed using a 128-pixel wide Gaussian boxcar. Even this approach did not reduce the residuals in regions A, B, and C. We also tried higher-order solutions but these did not improve the residuals either. It is likely that the increased residuals in these problematic regions can

only truly be addressed by deriving a new pixel-to-pixel flat (P-flat). Creating P-flats requires a large volume of data in order to reach the required S/N however, and we do not anticipate obtaining enough such data until at least the end of JWST's Cycle 1. Furthermore, only a subset of the full filter complement will be imaged as part of these P-flat observations. As such, the L-flats derived in NIS-011a will serve as the only flat field updates for some time, and for some filters they will be the only possible updates for several years.

For the short-wavelength (and more undersampled) filters, the RMS error is ≤ 0.0232 mags, and for the long-wavelength filters it is ≤ 0.0163 mags. We will continue to explore other fitting options in hopes of decreasing the residuals.

5 Acknowledgments

We wish to acknowledge the contributions, advice, and support from the following persons: Steph La Massa, René Doyon, Roberto Avila, and Jennifer Mack.

This work made use of Astropy: a community-developed core Python package and an ecosystem of tools and resources for astronomy. This work also made use of Photutils, an Astropy package for detection and photometry of astronomical sources (Bradley et al. 2022).

6 References

Astropy Collbarotation et al., 2022, ApJ, 935, 167

Bradley L., et al., 2022, astropy/photutils: 1.4.0, doi:10.5281/zenodo.6385735, https://zenodo.org/record/6385735

Cooper, R., Volk, K., NIRISS TR 7 Oct 2022, "NIRISS Commissioning Results: NIS-009"

JWST User Documentation (JDox). Baltimore, MD: Space Telescope Science Institute; 2016-2022 https://jwst-docs.stsci.edu

Mack, J., Sabbi, E., & Dahlen, T. 2013, WFC3 ISR 2013-10, "In-flight Corrections to the WFC3 UVIS Flat Fields"

Mack, J., Bohlin, R., Gilliland, R., van der Marel, R., Blakeslee, J., & DeMarchi, G. 2002, ACS ISR 2002-008, "ACS L-Flats for the WFC"

Martel, A., Sohn, T., Goudfrooij, P., Vila, M., B., & Zhou, J., NIRISS TR 15 Feb 2020, "NIRISS External Flat Fields (L-flats) and Distortion"

van der Marel, R. 2003, ACS ISR 2003-10, "Determination of Low-Frequency Flat-Field Structure from Photometry of Stellar Fields"