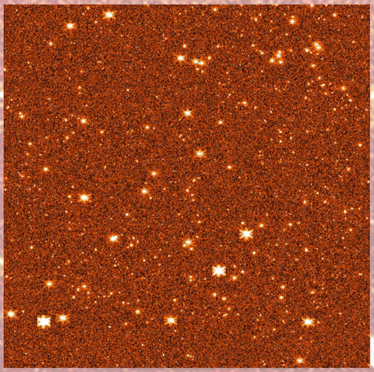


JWST

User Documentation for Cycle 1:

Data Calibration and Analysis

This PDF was last updated on February 12, 2018.
For the most current information, visit
https://jwst-docs.stsci.edu
and use on-line documentation.

JWST Calibration Programs and Data

The JWST calibration program will obtain the necessary reference files used by the pipeline to remove instrumental signatures and produce data products that are absolutely calibrated.

Introduction

The overall goal of the JWST pipeline is to produce data products that minimize instrumental signatures and are absolutely calibrated in physical units. This is done through the use of extensive ground-test data and on-orbit observations.

Absolute calibration

The requirements for the JWST absolute flux, wavelength, and position calibration are given in the tables below. At this point, we expect to meet all these requirements. The planned calibration programs are sufficient to meet or exceed these requirements.

Flux

The requirements for absolute flux calibration vary by instrument and observation mode. They are listed in Table 1. For the details of targets, etc. see the Absolute Flux Calibration article. Note that the needed flux prediction accuracy on the targets exceeds those listed in the table as there are other terms involved beyond the prediction accuracy (aperture correction accuracy, stability, etc.).

Table 1. Absolute flux requirements

Instrument	Imaging	Coronagraphy	Spectroscopy[1]
NIRCam	5%	5%	10%[2]
NIRSpec	N/A	N/A	10%
NIRISS	5%	N/A	10%[2]
MIRI	5%	15%	15%

[1] For sources observed through slits, the value is for well centered observations.

[2] Not a formal requirement, by analogy with NIRSpec spectroscopy.

Wavelength

The requirements for absolute wavelength calibration vary by instrument. They are listed in Table 2. For the details of targets, etc. see the Absolute Wavelength Calibration article. Note that the needed wavelength prediction accuracy on the targets exceeds those listed in the table as there are other terms involved beyond the prediction accuracy (aperture centering accuracy, stability, etc.).

Table 2. Absolute wavelength requirements

Instrument	Accuracy (resolution element)
NIRCam	10%[1]
NIRSpec	12.5%
NIRISS	10%
MIRI	10%

[1] Not a formal requirement, by analogy with NIRISS spectroscopy.

Position (astrometry)

The requirement for absolute position (astrometric) calibration is 5 mas. Specifically, the 5 mas refers to the uncertainty in the observatory field distortion correction within any instrument and the guider. For details see the article titled Absolute Astrometric Calibration.

Cycle 1 calibration programs

The instrument cycle 1 calibration programs will provide a set of deliverables needed to remove instrumental signatures and calibrate the JWST observations. Revisions to these programs are in progress with updates expected near the end of 2017.

The main deliverables from the calibration program are reference files for the pipeline.

The cycle 1 calibration program will be used in combination with any similar observations taken during the JWST commissioning to produce the highest quality reference files.

Example Deliverables

- darks
- bad pixel maps
- read noise measurements
- flat fields (both pixel and illumination versions)
- linearity corrections
- persistence characterization
- PSF images and characterization
- astrometric solutions (including distortion characterization)
- wavelength calibration
- line spread function characterization (for the spectrographs)
- absolute (spectro)photometric calibration
- relative throughput of NIRSpec apertures
- NIRSpec MSA shutter contrast and failed shutter maps
- NIRISS WFSS spectral traces

Absolute Astrometric Calibration

Astrometric calibration is required for efficient operations and to determine the absolute coordinates of observed sources in the sky. It relies on determining the mutual positions of the instrument and guiding cameras in the focal plane, on measuring the distortions of the instruments, and on the use of an accurate guide star catalog.

Focal plane alignment

The target acquisition and placement processes (see details here) rely on knowing the locations of the various instruments relative to the guiders. This information is derived from calibration observations and carried in the Science Instrument Aperture Files (SIAF, Cox et al. 2009, Cox & Lallo 2017 not yet released). The SIAF files contain parameters that provide coordinates transforms to and from four different frames: (1) the detector frame, (2) the science frame, which is an offset, flipped, and/or trimmed version of the detector frame, (3) the distortion corrected "ideal" frame, and (4) the v2/v3 plane, a spherical coordinate system that relates all instruments to the observatory.

Distortion calibration

An accurate distortion solution is critical for several JWST observing modes, e.g. NIRSpec multi object spectroscopy. In addition, for most observations dithers will be required to improve the PSF sampling, remove detector effects, and mitigate the impact of cosmic rays. A reliable dithers execution calls for good knowledge of the distortion in both the fine guidance sensor and the science instrument. The requirement is that the uncertainty on the field distortion determination for any instrument and guider is smaller than 5 milli-arcsecond (mas) RMS per axis.

Astrometric calibration plan

The JWST astrometric calibration will be performed by observing a field in the Large Magellanic Cloud. The focal plane alignment will be determined by imaging many stars with accurately known positions in both the guiders and the science instruments. The distortion calibration will consist in fine-guiding off of a star in one of the guiders while taking images of the calibration field with the respective science instrument. A source finding algorithm will locate the sources in these images and the stars' centroids will be measured. The astrometric comparison with the reference catalog yields the distortion solution.

The absolute astrometric accuracy of mapping pixels to the sky for JWST images reduced with the default JWST pipeline is limited by the absolute pointing accuracy at a level of approximately 0.5".

JWST astrometric calibration field

The JWST Astrometric Reference Field is a 5'×5' area in the LMC and located in the southern continuous-viewing zone. The field is characterized by a relatively homogeneous distribution of stars well-suited for JWST calibration, thus facilitating the derivation of the distortion solution in all JWST's instruments (Anderson 2008; Anderson 2016). In 2006 this field was observed with ACS on HST, providing a catalog of more than 200,000 isolated stars with positional accuracies of about 1 mas.

Incorporation of Gaia astrometry

The absolute astrometric reference frame of JWST will be based on the all-sky catalogs produced by the ESA Gaia mission. The astrometric catalogs for focal plane alignment and distortion calibration will be anchored on the Gaia reference frame, and the same is planned for the guide star catalog, thus leveraging the Gaia precision for JWST. For instance, Sahlmann (2017) investigated the astrometric agreement of common sources between the JWST calibration field catalog as derived from the ACS observations and the first data release of the Gaia mission. The offsets, rotation, scale differences, skew, and higher-order geometric distortions between both catalogs were determined and, under the assumption that Gaia data are accurate and distortion-free, a correction to the JWST calibration field catalog was derived.

References

Diaz-Miller 2008
Anderson 2008, Ref: JWST-STScI-001378
Cox & Lallo 2017, Ref: JWST-STScI-001550
Anderson 2016, Ref: JWST-STScI-005361
Sahlmann 2017, Ref: JWST-STScI-005492
JWST technical documents

Absolute Wavelength Calibration

The goal of the JWST absolute wavelength calibration is to transform spectra from detector pixel space to physical units of wavelength. To achieve this goal and meet wavelength accuracy requirements, on-orbit measurements with internal lamps and celestial calibrators will be used.

Introduction

The JWST science instruments offer a variety of spectroscopic capabilities. In the near-IR domain NIRCam provides grism slitess spectroscopy and NIRISS offers wide field slitless spectroscopy and single object slitless spectroscopy.

NIRSpec has 4 different spectroscopic modes:

- multi-object spectroscopy (MOS) with the micro-shutter assembly (MSA),
- 3-D imaging spectroscopy with the integral field unit (IFU),
- high-contrast single object spectroscopy with the fixed slits (FSs), and
- high throughput bright object time-series (BOTS) spectroscopy with the NIRSpec wide aperture.

In the mid-IR MIRI delivers low-resolution slitted and slitless spectroscopy and medium resolution integral field unit (IFU) spectroscopy. For calibration purposes, this wide range of observations requires coordination: the NIR instruments can use the same calibrators for many modes, and MIRI may need NIR data for planning purposes. Much effort is being invested now in defining the JWST wavelength calibrators list; this page will be updated when more detailed information is available.

Wavelength calibration requirements

Table 1. Accuracy requirements for each of the JWST science instruments

Science instrument	Requirement
MIRI	For all spectroscopic modes (MRS and LRS) the wavelengths shall be known to within 1/10 of a resolution element for an unresolved line with S/N = 50 after calibration.
NIRCam	No formal requirement
NIRISS	The wavelength solution for both WFSS grisms and orders 1 and 2 of the SOSS cross-dispersed grism shall be known to within 1/10th of a resolution element after calibration.
NIRSpec	For all spectroscopic modes the wavelength scale shall be determined with an accuracy of better than 1/8 of a resolution element after calibration

The requirement for the NIRISS grisms is taken to refer to a bright point source. For these slitless spectroscopy modes the wavelength resolution is degraded for extended sources, and it depends on the detailed size and shape of the object in that case.

NIRSpec calibration strategy

The NIRSpec wavelength calibration is applied to data using a parametric optical model of the instrument. The model provides a means of calculating the wavelength value at each pixel, given the disperser and aperture, by following the optical path of light from a celestial source through the instrument to the detector plane. The model was adopted in part to provide the pipeline with a flexible way of dealing with complex aspects of wavelength calibration over the NIRSpec field of view. One issue of particular importance is the fact that the grating wheel positioning is not strictly repeatable; this can result in dispersion shifts of up to one pixel in the detector plane between identical spectra taken before and after a grating wheel move. Any such shifts can be calculated with high accuracy using wheel tilt sensor telemetry, and a correction has been incorporated into the model calculations.

NIRSpec is the only science instrument that has on-board calibration sources suitable for wavelength characterisation. Initial calibration of the instrument model was carried out during ground testing using a combination of internal calibration lamps and an external Ar lamp.

The internal line lamps will be used to determine a final calibration of the instrument model on-orbit, as well as provide long-term monitoring. The lamps include two different line sources:

- "REF" lamp with erbium filter, which provides narrow absorption lines over a limited range of wavelengths.
- "LINE" lamps with interference filters that provide broad emission features over the full wavelength

range for all modes.

External observations of a celestial emission line source are still necessary in order to provide an independent check of the model.

Considerations for celestial calibrators

Ideally, a wavelength calibrator must comply with the following characteristics:

- Line density.
- Different source shapes: some spatially extended for MOS (not required but currently adopted) and IFU, and others point sources (grisms).
- The presence of companions should be taken considered to avoid confusion.
- Variable targets are not desired.

Feasibility of ground observations is also being considered: sources of K=15 need ~1hr from Keck/Gemini/VLT. Wavelengths longer than 3 μm are mostly out of reach.

Candidate targets

Planetary nebulae are well suited wavelength calibrators; they exhibit relatively narrow emission lines with moderate density, at least in the NIR. CLOUDY models can be used to extrapolate emission strengths at wavelengths non-accessible from the ground, for planning exposure times.

- Ground NIR observations are required to constrain the input physical parameters of the line models.
- It is also possible to use existing Spitzer observations for MIR.

Table 2. Information on some of the celestial calibrators already selected

Instrument	Calibrator	RA (J2000)	DEC (J2000)	Type	K magnitude
NIRSpec	NGC 6543	17:58:33	+66:38:00	Planetary Nebulae	8.34
NIRISS	SMP LMC 58	05:24:20.81	-70:05:01.9	Planetary Nebulae	14.5
NIRCam	SMP LMC 58	05:24:20.81	-70:05:01.9	Planetary Nebulae	14.5
MIRI	TBD				

NIRSpec target: NGC6543

The “Cat’s Eye” planetary nebulae is located in the northern CVZ, its halo is well-matched to the NIRSpec MSA FOV, and optical observations suggest “inert” kinematics (ΔVr < 5 km/s, FWHM ~ 10 km/s). IR spectroscopy is available for the of core only. A proposal has been submitted to Keck (NASA time); the main results are:

- MOSFIRE H/K band observations have shown that halo knots are pure H_2 emission. Since CLOUDY is less applicable in this case, we have used separate shock models instead to estimate line strengths across the NIRSpec wavelength range.
- Enough density of lines for RV analysis for extrapolation to 3–5 µm. RVs are consistent with optical measurements, and the uncertainties & dispersion are well within wavelength calibration accuracy requirement for R=2700 (~14 km/s).

NIRCam/NIRISS target: SMP LMC 58

This nearly unresolved (~0.08”) planetary nebulae emission line source is located in the CVZ. It's K magnitude is 14.5, rendering it too faint for SOSS calibration. It has at least 3–5 lines per filter. In this case there is need ground-based NIR spectrum to verify line fluxes. We have obtained ground-based NIR spectroscopy to measure line fluxes, the data analysis is still in progress.

MIRI (and NIRISS SOSS) targets:

Be stars are active objects that a have suitable line/continuum ratios. They also exhibit good line density across large wavelength range, lines are not too bright and they are unresolved (SOSS, LRS) . However, they may be variable and are unresolved; there is need mapping for MRS observations. We are currently vetting a small list of Be targets in the LMC using ground-based NIR spectroscopy.

References

Stanghellini, L., et al. 2002, ApJ 575 178S
Optical Slitless Spectroscopy of Large Magellanic Cloud Planetary Nebulae: A Study of the Emission Lines and Morphology

JWST technical documents

Absolute Flux Calibration

The prime objective of the JWST flux calibration plan is to enable high accuracy relative and absolute flux calibration for all science instruments, by converting science data from instrumental to physical units. The overall plan is described, including the list of the stellar flux calibration sources.

Introduction

The JWST has a suite of four scientific instruments covering near- and mid-IR wavelengths. From about 0.6 to 5.3 µm, NIRCam, NIRISS, and NIRSpec provide a variety of imaging, coronagraphic and spectroscopic modes. From 5 to 28.5 µm, MIRI offers the same array of observing capabilities. For all these instruments/modes, the nominal requirement is 2% absolute flux prediction accuracy of standard stars, with the goal of improving it as much as possible. This is the required prediction accuracy to achieve the overall flux calibration requirements as these include other terms beyond the flux prediction accuracy (aperture correction accuracy, stability, etc.). To achieve this, a sample of stars will well known absolute flux and spectral shape across the 0.6 to 28.5 microns range will be used. Having this unified program will not only enable the cross-instrument calibration of the JWST science instruments, but also with HST, Spitzer and other ground-based telescopes. Details on how the JWST pipeline applies the photometric calibration are given here.

JWST science instruments calibration goals

To achieve the 2% goal (and potentially push below 1% for the predicted fluxes), multiple calibrator stars of the same spectral class are needed to account for the differences between the actual stars and the adopted stellar model atmospheres. Calibrators of different spectral types are required to control for systematic uncertainties in the stellar atmospheres modelling.

The JWST flux calibration sample is composed by hot stars (white dwarfs and OB stars, see Table 1), A-stars (Table 2), and G-stars (solar-like, see Table 3). These spectral types can be modeled to high accuracy. Previous works on Hubble calibrators (Bohlin & Cohen 2008, Bohlin 2010) have shown that there is random modeling noise on the order of 2% for an individual calibration star. Observing 4 stars reduces the random uncertainty to 1% for the average JWST flux calibrations in each spectral bin. The sample covers the sensitivity range for all instruments, and at least 5 of them is observable with each filter/grating (Gordon et al. 2009; Gordon & Bohlin 2012).

Sample of JWST primary standards

The sample of standard is listed in Tables 1 to 3 and these tables are updated versions of those given by Gordon & Bohlin 2012. The sample contains sources faint enough to be observed with NIRCam, as well as calibrators suitable to characterize the longer MIRI wavelengths. Stars with existing or planned HST/Spitzer

observations were favored. Hubble STIS spectroscopy obtained during 2010-2013 time frame will provide the basis for fitting the model atmospheres that will predict fluxes at JWST wavelengths. Ground-based observations of each calibration star in the optical and near- infrared will be also acquired, to provide an independent prediction of the stellar atmosphere parameters (e.g., T(eff) & log(g)). This sample should provide sufficient high quality calibration stars to meet the required absolute flux calibration JWST error budget for the imagers (NIRCam & MIRI). The predicted spectra for a representative set of the primary calibrators are shown in Figure 1 from 0.8 to 28 µm. The predicted spectra for all the calibrators stars are available from CALSPEC.
Note the current sample will be enhanced to provide sources for the calibration characterization for Cycle 1, not just the base flux calibration.

Table 1. JWST primary calibrators: hot stars

Name	RA (J2000)	Dec (J2000)	Spec. Type	B	V
ksi02 Cet	02 28 09.54	+08 27 36.2	B9III	4.28	4.39
lam Lep	05 19 34.52	-13 10 36.4	B0.5IV	4.28	5.00
10 Lac	22 39 15.68	+39 03 01.0	O9V	4.88	5.50
mu Col	05 45 59.89	-32 18 23.2	O9.5V	5.17	5.99
G191B2B	05 05 30.62	+52 49 54.0	DA0	11.781	12.764
GD71	05 52 27.51	+15 53 16.6	DA1	13.032	14.115
GD153	12 57 02.37	+22 01 56.0	DA1	13.346	14.308
LDS749B	21 32 16.01	+00 15 14.3	DBQ4	14.73	15.217
WD1057+719	11 00 34.31	+71 38 03.3	DA1.2	14.8	
WD1657+343	16 58 51.10	+34 18 54.3	DA1	16.1	

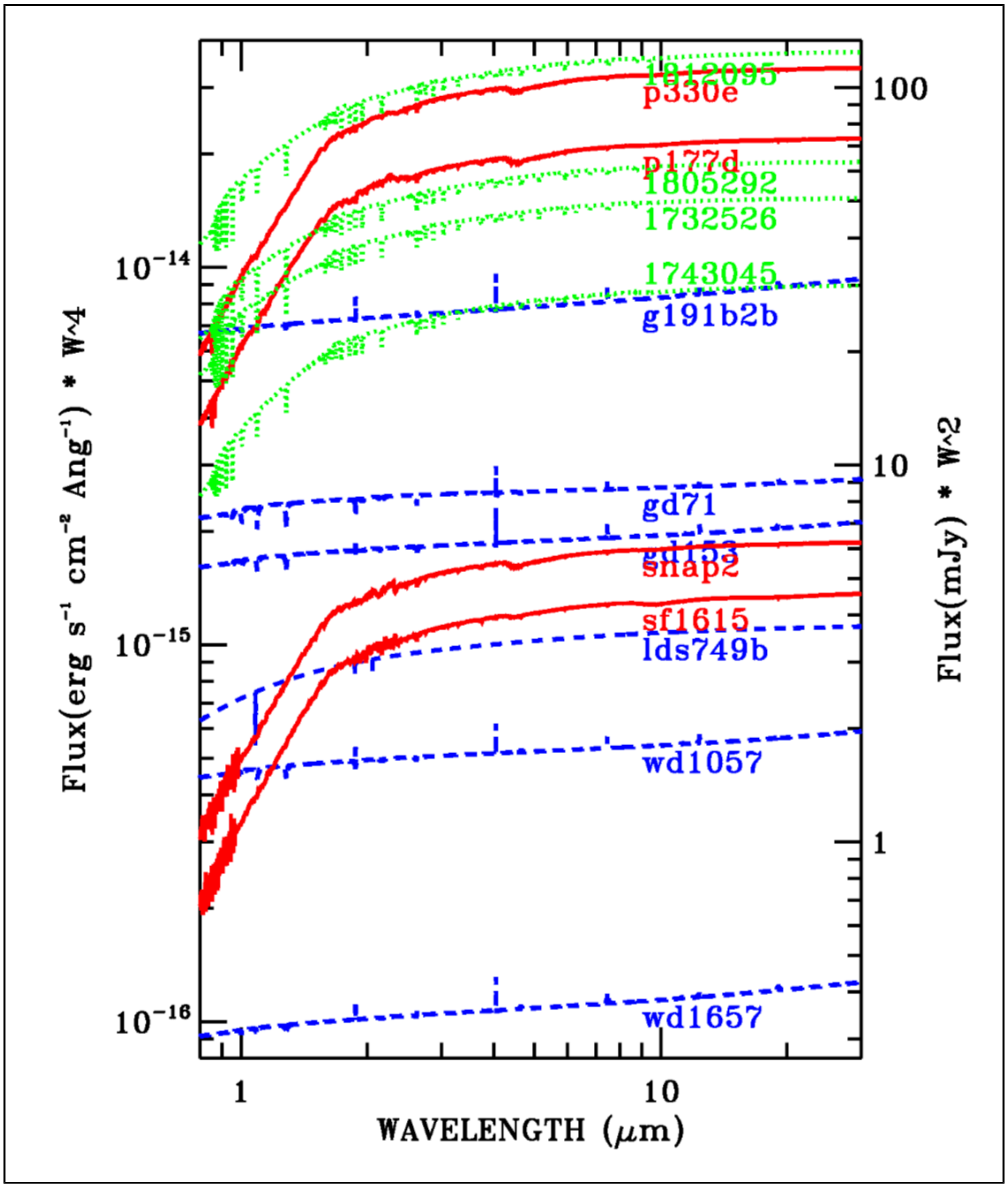
Table 2. JWST primary calibrators: A stars

Name	RA (J2000)	Dec (J2000)	SpType	V	K
HD 158485	17 26 04.84	+58 39 06.8	A4V	6.50	6.14
HD 14943	02 22 54.67	-51 05 31.7	A5V	5.91	5.44
HD 37725	05 41 54.37	+29 17 50.9	A3V	8.35	7.90
HD 163466	17 52 25.37	+60 23 46.9	A2	6.86	6.34
HD 116405	13 22 45.12	+44 42 53.9	A0V	8.34	8.48
HD 180609	19 12 47.20	+64 10 37.2	A0V	9.41	9.12
BD+60 1753	17 24 52.27	+60 25 50.7	A1V	9.67	9.64
1757132	17 57 13.25	+67 03 40.9	A3V	12.0	11.16
1812095	18 12 9.56	+63 29 42.3	A2V	11.736	11.286
1808347	18 08 34.75	+69 27 28.7	A3V	11.9	11.53
1802271	18 02 27.17	+60 43 35.6	A3V	11.985	11.832
1805292	18 05 29.3	+64 27 52.1	A1V	12.278	12.005
1732526	17 32 52.64	+71 04 43.1	A3V	12.530	12.254
1743045	17 43 04.48	+66 55 01.6	A5V	13.5	12.772

Table 3. JWST Primary calibrators: G stars

Name	RA (J2000)	Dec (J2000)	SpType	V	K
HD 159222	17 32 00.99	+34 16 16.1	G1V	6.56	5.00
HD 205905	21 39 10.15	-27 18 23.7	G2V	6.74	5.32
HD 106252	12 13 29.51	+10 02 29.9	G0	7.36	5.93
HD 37962	05 40 51.97	-31 21 04.0	G2V	7.85	6.27
HD 38949	05 48 20.06	-24 27 49.9	G1V	8.0	6.44
P330E	16 31 33.85	+30 08 47.1	G0V	13.01	11.379
P177D	15 59 13.59	+47 36 41.8	G0V	13.48	11.857
C26202	03 32 32.88	-27 51 48.0	G0-5	16.64	
SF1615+001A	16 18 14.23	+00 00 08.4	G0-5	16.75	
SNAP-2	16 19 46.13	+55 34 17.7	G0-5	16.2	

Figure 1. Representative spectra



Spectra of a representative sample of the proposed primary calibration sources: white dwarfs (blue, dashed), A-stars (green, dotted),and G-stars (red, solid). The multipliers, W4 and W2, on the ordinates are in units of m (Gordon & Bohlin 2012)

SED details

Pure hydrogen WDs are straightforward to model. GD153, GD71, and G191B2B are hydrogen WDs and have temperatures and gravities derived from fitting the models to their observed Balmer lines. Their models are normalized to precision Landolt V band photometry and are the primary absolute flux standards for all of the HST flux calibrations (Bohlin et al. 1995). For the cases of the pure helium WD LDS749B, the A stars, and the

G stars, their spectral distributions below 2.5 μm are measured from calibrated STIS and NICMOS spectrophotometry; and then the best fitting model is used to estimate the fluxes longward of 2.5 μm. However, NICMOS is no longer available; and STIS covers only wavelengths below 1 μm. Because the stellar models are most uncertain in regions of heavy line blanketing, broadband averages are used to find the best models, which match the observed fluxes to an rms scatter in the broad bands of less than 1% for all of our WD, A, and G standard stars. Thus, the continuum regions of the model extensions above 2.5 μm should be good to the same 1–2% that is the quoted precision for STIS+NICMOS.

Instrument sensitivities

Information about the JWST science instruments sensitivities is provided in the following pages for NIRCam, NIRISS, NIRSpec, and MIRI. The proposed list of calibrators has been defined to optimally cover these sensitivity ranges for the different instrument modes, based on the following sensitivity levels:
MAX observable flux = flux that can be observed in the normal observing modes (full frame and subarrays) without reaching saturation
MIN observable flux = flux that can be observed with a S/N of 200/50 in 3600 s for imaging/spectroscopy.

References

Bohlin et al. 1995, White Dwarf Standard Stars: G191-B2B, GD 71, GD 153, HZ 43, AJ, 110, 1316

Bohlin & Cohen 2008, NICMOS Spectrophotometry and Models for A Stars, AJ, 136, 1171

Bohlin 2010, Hubble Space Telescope Spectrophotometry and Models for Solar Analogs, AJ, 139, 1515

Gordon et al. 2009, JWST Absolute Flux Calibration I: Proposed Primary Calibrators, JWST-STScI-001855

Gordon & Bohlin 2012, JWST Absolute Flux Calibration II: Expanded Sample of Primary Calibrators, JWST-STScI-002540

JWST technical documents

All Calibration Reference Files

The STScI/JWST pipeline automatically processes data for all available observing modes. Details of the pipeline architecture and algorithms can be found here. Each branch of the pipeline uses a set of instrument-specific reference files that ensure the pipeline meets its accuracy requirements.
The JWST pipeline provides science-ready calibrated data for all offered instruments/observing modes. The pipeline relies on the use of a set of reference files to ensure the correct and accurate process of the data. The reference files are instrument-specific, and are periodically updated as the data process evolves and the understanding of the instruments improves. The following links provide information on the individual reference files used by the JWST instrument in the different pipeline stages.

- MIRI specific reference files.
- NIRCam specific reference files
- NIRISS specific reference files
- NIRSpec specific reference files

MIRI Calibration Reference Files

This page provides information on the reference files used by the JWST pipeline to process JWST MIRI data. The information is organised by pipeline stages and lists all reference files needed for each step. Detailed information on the pipeline steps and files format is provided in the specified links.

Introduction

The JWST pipeline provides science-ready calibrated data for all MIRI observing modes: imaging, coronagraphic imaging, medium resolution spectroscopy (via an integral field unit), and low resolution spectroscopy (slitted and slitless). The pipeline relies on the use of a set of reference files essential to ensure the correct and accurate process of the data. This page gives details on the MIRI-specific reference files which are periodically updated. The following tables are organised according to the different pipeline stages, and provide information on all reference files used for the data reduction. In all cases xxxx replaces the version number of the file.

CALWEBB_DETECTOR1 reference files

This pipeline stage process the LVL1 up-the-ramp raw files and returns the level 2a files in units of (counts/sec). The pipeline steps flow for MIRI is different from that of the NIR instruments, and has been adjusted to better optimize the data treatment of the Si:As detectors in MIRI. The pipeline steps made use of the following reference files:

Table 1. MIRI reference files used by the JWST pipeline module calwebb_detector1.

Pipeline step	Reference files used	Usability notes
Data Quality Initialization	Bad pixel mask file: jwst_miri_mask_xxxx.fits	
Saturation check	Saturation file: jwst_miri_saturation_xxxx.fits	
Intrapixel Capacitance Correction	IPC file: jwst_miri_ipc_xxxx.fits	This step is included in the pipeline, but disabled by default. Pipeline step not applied to TSO data

Linearity Correction	Linearity file: jwst_miri_linearity_xxxx.fits	
First Frame Correction	Pipeline step not currently used.	Pipeline step not applied to TSO data
MIR RSCD	RSCD file: jwst_miri_rscd_xxxx.fits	
Last Frame correction	Step does not use any reference file.	Pipeline step not applied to in TSO data
Dark Subtraction	Dark file: jwst_miri_dark_xxxx.fits	
Reference Pixels Correction	Pipeline step does not use any reference file. The correction uses reference pixels information stored in each exposure.	
Persistence correction	This step uses three different files, that are currently being defined. This information will be updated soon. The traps decay functions separated by family types The trap density map. The persistence saturation map.	Pipeline step not applied to TSO data
Jump Detection	Gain file: jwst_miri_gain_xxxx.fits Readnoise file: jwst_miri_readnoise_xxxx.fits	
Ramp Fitting	Gain file: jwst_miri_gain_xxxx.fits Readnoise file: jwst_miri_readnoise_xxxx.fits	

CALWEBB_Dark

This pipeline branch is intended for use with dark exposures. It applies all of the same detector-level correction steps as CALWEBB_Sloper and uses the same reference files, but stops just before the application of the dark subtraction step.

CALWEBB_IMAGE2 reference files

In this pipeline stage individual exposures from the imaging modes are processed from counts/sec to absolute flux units. During Cycle 1 no MIRI imager observations will be flagged as TSO. Hence, if obtained, they will follow the standard imaging pipeline processing. Details on the pipeline steps flow are given here.

Table 2. MIRI reference files used by the JWST pipeline module calwebb_image2.

Pipeline step	Reference files used	Usability notes
Assign WCS Information	Distortion file: jwst_miri_distortion_xxx.asdf Filter offset: jwst_miri_filteroffset_xxxx.asdf	
Background subtraction	Step does not use any reference file.	This pipeline step is not applied to Target Acquisition images. It subtracts a dedicated background image if requested by the oberver.
Flat Field Correction	Flat field file: jwst_miri_flat_xxxx.fits	
Flux Calibration	Photom file: jwst_miri_photom_xxxx.fits Area file: jwst_miri_area_xxxx.fits	Not applied to target acquisition images.
Rectify 2D Image		

CALWEBB_SPEC2 reference files

This step process Individual exposures from the spectroscopic modes from counts/sec to absolute flux and wavelength units. Nominally, this stage works on individual exposures. This page gives details on the pipeline flow.

Table 3. MIRI reference files used by the JWST pipeline module calwebb_spec2.

Pipeline step	Reference files used	Usability notes
Assign WCS Information	Distortion file: jwst_miri_distortion_xxx.asdf Spectroscopic WCS: jwst_miri_specwcs_xxxx.asdf jwst_miri_specwcs_xxxx.fits Regions: jwst_miri_regions_xxxx.asdf V2V3: jwst_miri_v2v3_xxxx.asdf Wavelength Range: jwst_miri_wavelengthrange_xxxx.asdf	jwst_miri_specwcs_xxxx.asdf and V2V3 needed for MRS jwst_miri_specwcs_xxxx.fits needed for LRS
Background subtraction	Step does not use any reference file.	Step only applied to the MIRI LRS
Flat Field Correction	Flat field: jwst_miri_flat_xxxx.fits	
Point vs Extended Decision	Step does not use any reference file.	
Source Type Determination	Step does not use any reference file.	
Stray Light Correction	Stray light mask: jwst_miri_straymask_xxxx.fits	Only applied to MIRI MRS data
Fringe Correction	jwst_miri_fringe_xxxx.fits	Only applied to MIRI MRS data
Pathloss correction	Ref. file currently not in CRDS	Only applied to MIRI LRS data
Flux Calibration	Photom: jwst_miri_photom_xxxx.fits RELSENS RELSENS2D	RELSENS used for LRS data RELSENS2D used for MRS data
Cube Building	No reference files are read by this step. All the reference file selection and reading is done in the assign_wcs step.	
Extract 1D spectrum	jwst_miri_extract1d_xxxx.json	

CALWEBB_IMAGE3 reference files

Multiple exposures from the direct imaging modes (e.g., dither pattern or mosaic) are combined into a single rectified (distortion corrected) image. Details on the pipeline steps flow are given in this page.

Table 4. MIRI reference files used by the JWST pipeline module calwebb_image3.

Pipeline step	Reference files used	Usability notes
Tweakreg Catalog	Step does not use any reference file.	
Tweakreg	Step does not use any reference file.	
Skymatch	Step does not use any reference file.	
Outlier_detection	Step does not use any reference file.	
Resample	Step does not use any reference file.	
Source_catalog	Step does not use any reference file.	

CALWEBB_CORON3

This pipeline stage processes and combines multiple coronagraphic exposures.

Table 5. MIRI reference files used by the JWST pipeline module calwebb_coron3.

Pipeline step	Reference files used	Usability notes
Assemble Ref PSFs	Step does not use any reference file, but reference PSFs.	
Align Ref PSFs	Uses PSF weighting Ref. File	
Ref PSF Subtraction	Step does not use any reference file.	
Outlier Detection	Step does not use any reference file.	
Create Exp Level Products	Step does not use any reference file.	
Image Combination/Resample	jwst_miri_drizpars_xxxx.fits	

CALWEBB_SPEC3

Multiple individual exposures for the spectroscopic modes (i.e. dither patterns or mosaics) are combined into a single rectified cube and/or extracted spectrum.

Table 6. MIRI reference files used by the JWST pipeline module calwebb_spec3.

Pipeline step	Reference files used	Usability notes
Background Matching	Step does not use any reference file.	MRS only
Outlier Detection	Step does not use any reference file.	
Create Exp Level Products	Step does not use any reference file.	
Cube Creation	Step does not use any reference file.	
Spectral Extraction 1D	jwst_miri_extract1d_xxxx.json	
Spectral Leak Sub.	Need reference file (TBD)	

CALWEBB_TSO3

This pipeline stage processes individual exposures taken in the Time Series Observation (TSO) modes.

Table 7. MIRI reference files used by the JWST pipeline module calwebb_tso3.

Pipeline step	Reference files used	Usability notes
Outlier Detection	Step does not use any reference file.	
Create Exp Level Products	Step does not use any reference file.	
Extract Photometry	Ref. file TBD	MIRI Imaging only.
Extract 1D Spectrum	jwst_miri_extract1d_xxxx.json	MIRI Spectroscopy only.
White-Light Photometry	Step does not use any reference file.	MIRI Spectroscopy only.

NIRCam Calibration Reference Files

This page provides information on the reference files used by the JWST pipeline to process NIRCam data. The information is presented by pipeline stages. Details on the pipeline flow and steps, files format and their usage are provided in the specified links.

Introduction

The JWST NIRCam instrument offers 5 science observing modes that are automatically processed by the JWST pipeline: imaging, coronagraphic imaging, wide field slitless spectroscopy, time series imaging and grism time series. The following sections provide information on the NIRCam-specific reference files used by the pipeline, organised according to the different pipeline stages. In all cases xxxx replaces the version number of the file.

CALWEBB_DETECTOR1 reference files

This pipeline stage process the LVL1 up-the-ramp raw files and returns the level 2a files in units of (counts/sec). To achieve that several the diverse steps in the correspondent NIR pipeline flow use the following reference files:

Table 1. NIRCam reference files used by the JWST pipeline module calwebb_detector1.

CALDETECTOR1 pipeline step	Reference file used	Usability notes
Data Quality Initialization	jwst_nircam_mask_xxxx.fits One reference files per each detector, Full and subarray modes	
Saturation check	jwst_nircam_saturation_xxxx.fits	
Interpixel Capacitance Deconvolution	jwst_nircam_ipc_xxxx.fits	Step included in the pipeline, but disabled by default. Pipeline step not applied to time series imaging data
Superbias Subtraction	jwst_nircam_superbias_xxxx.fits	
Reference Pixels Correction	N/A. The correction uses reference pixels information stored in each exposure	
Linearity Correction	jwst_nircam_linearity_xxxx.fits	
Persistence correction	Persat: jwst_nircam_persat_xxxx.fits Trap density: jwst_nircam_trapdensity_xxxx.fits Trap pars: jwst_nircam_trappars_xxxx.fits	Pipeline step not applied to time series imaging data
Dark Subtraction	jwst_nircam_dark_xxxx.fits	
Jump Detection	This pipeline step uses gain and readnoise information Gain: jwst_nircam_gain_xxxx.fits Read noise: jwst_nircam_readnoise_xxxx.fits	
Ramp Fit	Gain: jwst_nircam_gain_xxxx.fits Read noise: jwst_nircam_readnoise_xxxx.fits	

CALWEBB_Dark

This pipeline branch is intended for use with dark exposures. It applies all of the same detector-level correction steps as CALWEBB_Slopes and uses the same refrence files, but stops just before the application of the dark subtraction step.

CALWEBB_IMAGE2 reference files

In this pipeline stage individual exposures from the imaging modes are processed from counts/sec to absolute flux units. Details on the pipeline steps flow are given here.

Table 2. NIRCam reference files used by the JWST pipeline module calwebb_image2.

Pipeline step	Reference file used	Usability notes
Assign WCS Information	Distortion file: jwst_nircam_distortion_xxxx.asdf	
Background Subtraction	Step does not use any reference file.	Step applies to science data if the user requests a dedicated background image. Not applied to Target Acquisition images.
Flat Field Correction	Flat field: jwst_nircam_flat_0105.fits	
Flux Calibration	Photom: jwst_nircam_photom_0031.fits Area: jwst_nircam_area_0001.fits	Not applied to Target Acquisition images.
Rectify 2D image	Step does not use any reference file.	

CALWEBB_SPEC2 reference files

This pipeline stage process Individual exposures from the spectroscopic modes from counts/sec to absolute flux and wavelength units. Nominally, this stage works on individual exposures obtained for the NIRCam wide field slitless grism observations, but the steps are not yet implemented.

Table 3. NIRCam reference files used by the JWST pipeline module calwebb_spec2.

Pipeline step	Reference file used	Usability notes
Assign WCS Information	Distortion file: jwst_nircam_distortion_xxxx.asdf Specwcs: File not available in CRDS	
Background Subtraction	Step does not use any reference file.	
Subwindow extraction	Step does not use any reference file.	
Flat Field	Flat filed: jwst_nircam_flat_xxxx.fits	
Point vs Extended Decision	Step does not require any reference file	
Photometric Correction	Photom: jwst_nircam_photom_xxxx.fits Area: jwst_nircam_area_xxxx.fits	
Rectified 2D 3D product		

CALWEBB_IMAGE3 reference files

Multiple exposures from the direct imaging modes (e.g., dither pattern or mosaic) are combined into a single rectified (distortion corrected) image. Details on the pipeline steps flow are given in this page.

Table 4. NIRCam reference files used by the JWST pipeline module calwebb_image3.

Pipeline step	Reference file used	Usability notes
Tweakreg Catalog	Step does not use any reference file.	
Tweakreg	Step does not use any reference file.	
Skymatch	Step does not use any reference file.	
Outlier_detection	Step does not use any reference file.	
Resample	jwst_nircam_drizpars_xxxx.fits	
Source_catalog	Step does not use any reference file.	

CALWEBB_SPEC3

Multiple individual spectroscopic exposures (i.e. dither patters or mosaics) are combined into a single rectified cube and/or extracted spectrum.

Table 5. NIRCam reference files used by the JWST pipeline module calwebb_spec3.

Pipeline step	Reference file used	Usability notes
Outlier Detection	Step does not require any extra reference file.	
Create Exp Level Products		
Cube Creation	Step does not use any reference file.	
Spectral Extraction 1D		

CALWEBB_CORON3

This pipeline stage processes and combines multiple coronagraphic exposures.

Table 6. NIRCam reference files used by the JWST pipeline module calwebb_coron3.

Pipeline step	Reference file used	Usability notes
Assemble Ref PSFs	Step does not use any reference file, but reference PSFs.	
Align Ref PSFs	Step does not use any reference file.	
Background Matching		Not part of the baseline pipeline
Ref PSF Subtraction	Step does not use any reference file.	
Outlier Detection	Step does not use any reference file.	
Image Combination	Step does not use any reference file.	
Create Exp Level Products	Step does not use any reference file.	
Ancillary Science Products		Not part of the baseline pipeline

NIRISS Calibration Reference Files

This page provides information on the reference files used by the JWST pipeline to process NIRISS data. The information is sorted by pipeline stages. Detailed information on the pipeline steps and files format is provided in the specified links.

Introduction

The JWST NIRISS instrument offers 4 science observing modes that are automatically processed by the JWST pipeline: wide field slitless spectroscopy, single object slitless spectroscopy, aperture masking interferometry, and imaging. The following sections provide information on the NIRISS-specific reference files used by the pipeline, organised according to the different pipeline stages.

CALWEBB_DETECTOR1

This pipeline stage process the LVL1 up-the-ramp raw files and returns the level 2a files in units of (counts/sec). All NIRISS data goes though the entire set of pipeline steps applied in CALDETECTOR1. To achieve that several the diverse steps in the correspondent NIR pipeline flow use the following reference files:

Table 1. NIRISS reference files used by the JWST pipeline module calwebb_detector1.

Pipeline step	Reference files used	Usability notes
Data Quality Initialisation	jwst_niriss_mask_xxxx.fits	
Saturation check	jwst_niriss_saturations_xxxx.fits	
Interpixel Capacitance Correction	jwst_niriss_ipc_xxxx.fits	Step included in the pipeline, but disabled by default.
Superbias Subtraction	jwst_niriss_superbias_xxxx.fits One reference file per readout mode/subarray.	
Reference Pixels Correction	N/A. The correction uses reference pixels information stored in each exposure	
Linearity Correction	jwst_niriss_linearity_xxxx.fits	
Persistence Correction	Persat: jwst_niriss_persat_xxxx.fits Trap density: jwst_niriss_trapdensity_xxxx.fits Trap pars: jwst_niriss_trappars_xxxx.fits	
Dark Subtraction	jwst_niriss_dark_xxxx.fits One reference file per readout mode/subarray.	
Jump Detection	Gain: jwst_niriss_gain_xxxx.fits Read noise: jwst_niriss_readnoise_xxxx.fits	
Ramp fit	Gain: jwst_niriss_gain_xxxx.fits Read noise: jwst_niriss_readnoise_xxxx.fits	

CALWEBB_Dark

This pipeline branch is intended for use with dark exposures. It applies all of the same detector-level correction steps as CALWEBB_Slopes and uses the same reference files, but stops just before the application of the dark subtraction step.

CALWEBB_IMAGE2 reference files

In this pipeline stage individual exposures from the imaging modes are processed from counts/sec to absolute flux units. Details on the pipeline steps flow are given here.

Table 2. NIRISS reference files used by the JWST pipeline module calwebb_image2.

Pipeline step	Reference files used	Usability notes
Assign WCS Information	Distortion file: jwst_niriss_distortion_xxxx.asdf	
Background Subtraction	Step does not use any reference file.	Step applies to science data if the user requests a dedicated background image. Not applied to Target Acquisition images.
Flat Field Correction	Flat field: jwst_niriss_flat_xxxx.fits	
Flux Calibration	Photom: jwst_niriss_photom_xxxx.fits Area: jwst_niriss_area_xxxx.fits	Not applied to Target Acquisition images.
Rectify 2D image	Step does not use any reference file.	

CALWEBB_SPEC2 reference files

This pipeline stage process Individual exposures from the spectroscopic modes from counts/sec to absolute flux and wavelength units. Nominally, this stage works on individual exposures obtained for the NIRISS wide field slitless grism observations, but the steps are not yet implemented.

Table 3. NIRISS reference files used by the JWST pipeline module calwebb_spec2.

Pipeline step	Reference files used	Usability notes
Assign WCS Information	Distortion file: jwst_niriss_distortion_xxxx.asdf	
Background Subtraction	WFSS uses a reference file (To be defined). SOSS uses a reference file (master background).	
Subwindow extraction	Step does not use any reference file.	
extract_2d	jwst_niriss_wavecorr_xxxx.asdf	Step applied in NIRISS WFSS only
Flat Field	Flat file: jwst_niriss_flat_xxxx.fits	
Point vs Extended Decision	Step does not require any reference file	
Photometric Correction	Photom file: jwst_niriss_photom_xxxx.fits Area: jwst_niriss_area_xxxx.fits	
Rectified 2D 3D product		
Spectral Extraction 1d	jwst_niriss_extract1d_xxxx.json	

CALWEBB_IMAGE3 reference files

Multiple exposures from the direct imaging modes (e.g., dither pattern or mosaic) are combined into a single rectified (distortion corrected) image. Details on the pipeline steps flow are given in this page.

Table 4. NIRISS reference files used by the JWST pipeline module calwebb_image3.

Pipeline step	Reference files used	Usability notes
Tweakreg Catalog	Step does not use any reference file.	
Tweakreg	Step does not use any reference file.	
Skymatch	Step does not use any reference file.	
Outlier_detection	Step does not use any reference file.	
Resample	Step does not use any reference file.	
Source_catalog	Step does not use any reference file.	

CALWEBB_SPEC3

Multiple individual exposures for the spectroscopic modes (i.e. dither patters or mosaics) are combined into a single rectified cube and/or extracted spectrum.

Table 5. NIRISS reference files used by the JWST pipeline module calwebb_spec3.

Pipeline step	Reference files used	Usability notes
Outlier Detection	Step does not require any extra reference file.	
Create Exp Level Products		
Cube Creation	Step does not use any reference file.	
Spectral Extraction 1D	jwst_niriss_extract1d_xxxx.json	

CALWEBB_AMI3

The Level-3 Aperture Mask Interferometry pipeline is intended to be applied to associations of individual calibrated NIRISS AMI exposures and is used to compute fringe parameters and correct science target fringe parameters using observations of reference targets. Details on the individual pipeline steps can be found here.

Table 6. NIRISS reference files used by the JWST pipeline module calwebb_ami3.

Pipeline step	Reference files used	Usability notes
Analize	jwst_niriss_throughput_xxxx.fits	
Average	Step does not use any reference file.	
Normalize	Step does not use any reference file.	

NIRSpec Calibration Reference Files

This page provides information on the reference files used by the JWST pipeline to process NIRSpec data. The information is presented by pipeline stages. Details on the pipeline flow and steps, files format and their usage are accesible in the linked pages.

Introduction

The JWST NIRSpec instrument offers 4 science observing modes that are automatically processed by the JWST pipeline: multi-object spectroscopy with the micro-shutter array, integral field spectroscopy, fixed slits spectroscopy and bright object time-series (BOTS) spectroscopy. The following sections provide information on the NIRSpec-specific reference files used by the pipeline, organised according to the different pipeline stages. In all cases xxxx replaces the version number of the file.

CALWEBB_DETECTOR1 reference files

This pipeline stage process the LVL1 up-the-ramp raw files and returns the level 2a files in units of (counts/sec). All processing steps in CALDETECTOR1, with the exception of Group Scale Correction, are applied to every NIRSpec exposure.To achieve that several the diverse steps in the correspondent NIR pipeline flow use the following reference files:

Table 1. NIRSpec reference files used by the JWST pipeline module calwebb_detector1.

Pipeline step	Reference files used	Usability notes
Group Scale Processing	Step does not use any reference file.	Pipeline step applied only in data using NRS IRS2 readout mode.
Data Quality Initialization	Bad pixel mask: jwst_nirspec_mask_xxxx.fits	
Saturation check	Saturation mask: jwst_nirspec_saturation_xxxx.fits	
Interpixel Capacitance Deconvolution	jwst_nirspec_ipc_xxxx.fits	Step included in the pipeline, but disabled by default. Pipeline step not applied in BOTS observations
Superbias Subtraction	Bias file: jwst_nirspec_superbias_xxxx.fits	
Reference Pixels Correction	Ref. pixel file: jwst_nirspec_refpix_xxxx.fits	Used in IRS2 readout mode.
Linearity Correction	Lin corre file: jwst_nirspec_linearity_xxxx.fits	
Persistence correction	Persistence Saturation: jwst_nirspec_persat_xxxx.fits Traps Density: jwst_nirspec_trapdensity_xxxx.fits Trap Pars: jwst_nirspec_trappars_xxxx.fits	Pipeline step not applied in BOTS observations
Dark Subtraction	Dark file: jwst_nirspec_dark_xxxx.fits	
Jump Detection	Gain: jwst_nirspec_gain_xxxx.fits Readnoise: jwst_nirspec_readnoise_xxxx.fits	
Ramp Fit	Gain: jwst_nirspec_gain_xxxx.fits Readnoise: jwst_nirspec_readnoise_xxxx.fits	

CALWEBB_Dark

This pipeline branch is intended for use with dark exposures. It applies all of the same detector-level correction steps as CALWEBB_Slopes and uses the same refrence files, but stops just before the application of the dark subtraction step.

CALWEBB_IMAGE2 reference files

Although NIRSpec is primarily an spectrometer, there are several instances in which it is taking images (e.g. target acquisition, and confirmation images). These images also need to be processed by CALWEBB_IMAGE2 so the WCS information, needed later to apply distortion solution, is properly populated. For NIRSpec some of the standard calibration steps of this pipeline stage (e.g. flux calibration) are not applied. Details on the pipeline steps flow are given here.

Table 2. NIRSpec reference files used by the JWST pipeline module calwebb_image2.

Pipeline step	Reference files used	Usability notes
Assign WCS Information	Camera: jwst_nirspec_camera_xxxx.asdf Colimator: jwst_nirspec_collimator_xxxx.asdf Disperser: jwst_nirspec_disperser_xxxx.asdf Fore: jwst_nirspec_fore_xxxx.asdf FPA: jwst_nirspec_fpa_xxxx.asdf IFUfore: jwst_nirspec_ifufore_xxxx.asdf IFUpost: jwst_nirspec_ifupost_xxxx.asdf IFUslicer: jwst_nirspec_msa_xxxx.asdf MSA: jwst_nirspec_msa_xxxx.asdf OTE: jwst_nirspec_ote_xxxx.asdf Wavelength range: jwst_nirspec_wavelengthrange_xxxx.asdf Verify that these are all used by the imaging data	
Flat Field Correction	Dflat: jwst_nirspec_dflat_xxxx.fits Need to confirm reference file	
Rectify 2D image	Step does not use any reference file.	

CALWEBB_SPEC2 reference files

This step process Individual exposures from the spectroscopic modes from counts/sec to absolute flux and wavelength units. Nominally, this stage works on individual exposures. Details on the pipeline flow are given here.

Table 3. NIRSpec reference files used by the JWST pipeline module calwebb_spec2.

Pipeline step	Reference files used	Usability notes

Assign WCS Information	Camera: jwst_nirspec_camera_xxxx.asdf Colimator: jwst_nirspec_collimator_xxxx.asdf Disperser: jwst_nirspec_disperser_xxxx.asdf Fore: jwst_nirspec_fore_xxxx.asdf FPA: jwst_nirspec_fpa_xxxx.asdf IFUfore: jwst_nirspec_ifufore_xxxx.asdf IFUpost: jwst_nirspec_ifupost_xxxx.asdf IFUslicer: jwst_nirspec_msa_xxxx.asdf MSA: jwst_nirspec_msa_xxxx.asdf OTE: jwst_nirspec_ote_xxxx.asdf Wavelength range: jwst_nirspec_wavelengthrange_xxxx.asdf	
Background Subtraction	Step does not require any reference file.	
Imprint Subtraction	Step does not require any reference file.	Step applied to NIRSpec MOS and IFU data.
MSA failed open flagging	jwst_nirspec_msaoper_xxxx.json	Step applied to NIRSpec MOS and Fixed Slit data
Extract 2d spectra	jwst_nirspec_wavecorr_xxxx.asdf	Step applied to NIRSpec Fixed Slit and MOS data.
Flat Field	Sflat: jwst_nirspec_sflat_xxxx.fits Fflat: jwst_nirspec_fflat_xxxx.fits Dflat: jwst_nirspec_dflat_xxxx.fits	
Point vs Extended Decision	Step does not require any reference file	
Source Type Determination	Step does not require any reference file.	
Photometric Correction	Photom: jwst_nirspec_photom_xxxx.fits Area: jwst_nirspec_area_xxxx.fits	

Resample	Step does not require any reference file.	Step applied to NIRSpec Fixed Slit and MOS data.
Cube Build	No reference files are read in this step. The necessary information is already provided by the "Assign WCS information" step.	Step applied to NIRSpec IFU data.
Extract 1d spectra	jwst_nirspec_extract1d_xxxx.json	

CALWEBB_SPEC3

Multiple individual exposures for the spectroscopic modes (i.e. dither patters or mosaics) are combined into a single rectified cube and/or extracted spectrum.

Table 4. NIRSpec reference files used by the JWST pipeline module calwebb_spec3.

Pipeline step	Reference files used	Usability notes
Master Background Sub.	Step does not use any reference file.	
Outlier Detection	Step does not use any reference file.	
Create Exp Level Products	Step does not use any reference file.	
Cube Creation	Step does not use any reference file.	
Extract 1d spectra	jwst_nirspec_extract1d_xxxx.json	

CALWEBB_TSO3

This pipeline stage processes individual exposures taken in the Time Series Observation (TSO) modes.

Table 5. NIRSpec reference files used by the JWST pipeline module calwebb_tso3.

Pipeline step	Reference files used	Usability notes
Outlier Detection	Step does not use any reference file.	
TSO Background Subtraction	Step does not use any reference file.	Not used by baseline pipeline.
Create Exp Level Products	Step does not use any reference file.	
Extract 1D Spectrum	Step does not use any reference file.	
White-Light Photometry	Step does not use any reference file.	

JWST Data Reduction Pipeline

The JWST Data Reduction Pipeline is a Python software suite that automatically process the data taken by the JWST instruments NIRCam, NIRSpec, NIRISS, MIRI, and FGS to remove instrumental signatures from the observations.
The calibration pipeline processes data from all instruments and observing modes. It produces both fully calibrated individual exposures and high level data products (mosaics, extracted spectra, etc.).

The calibration pipeline is organized in stages by type of observations. The algorithms and overall structure of the pipeline are decided by an inclusive working group.

The calibration pipeline is coded in Python. The details of how to install the pipeline, run the pipeline including tutorials, and the detailed code documentation including the data structures used are provided.

Primer and Tutorials

This article presents a high-level summary of how users can run the various JWST calibration pipeline modules, giving examples of several types of common operations.

Introduction

The JWST calibration pipeline modules and individual calibration steps can be executed as command-line tasks or from within Python (either a simple python shell or something like a Jupyter notebook). There are 3 main stages of pipelines required to completely process a set of exposures for a given observation.

- Stage 1: Apply detector-level corrections to the raw data for individual exposures and produce count rate (slope) images from the "ramps" of non-destructive readouts
- Stage 2: Apply physical corrections (e.g. slit loss) and calibrations (e.g. absolute fluxes and wavelengths) to individual exposures
- Stage 3: Combine the fully calibrated data from multiple exposures

Some of these stages use different pipeline modules depending on the observation mode, as shown in Table 1. The sections below describe and provide examples for running some of these pipeline stages.

Table 1. Pipeline modules used in stages 1, 2, and 3

Stage	Mode	Pipeline
1	All	calwebb_detector1
2	Imaging	calwebb_image2
2	Spectroscopy	calwebb_spec2
3	Imaging	calwebb_image3
3	Spectroscopy	calwebb_spec3
3	Coronagraphy	calwebb_coron3
3	Time Series (imaging and spectroscopy)	calwebb_tso3
3	Aperture Masking Interferometry	calwebb_ami3

CALWEBB_DETECTOR1

The `calwebb_detector1` pipeline applies detector-level corrections to individual exposures, such as dark subtraction and linearity correction, and produces corrected countrate images as its final product. Complete details of the `calwebb_detector1` pipeline are available in the pipeline reference documents. Briefly, the `calwebb_detector1` pipeline takes raw data in the form of an `_uncal.fits` product as its input, which contains the original 4D data ramps of an exposure, applies a large list of detector-level corrections to the raw ramps, and then finally computes slopes from the ramps of each pixel to produce countrate images. The final output is a `_rate.fits` product. A `_rateints.fits` product is also created if the exposure contained multiple integrations. The `calwebb_detector1` pipeline is designed to be applied to raw data from all instruments and all observing modes (imaging, spectroscopy, coronagraphy, etc.).

The pipeline can be run on an individual exposure from the command line by using the `strun` task like this:

```
$ strun calwebb_detector1.cfg jw96090001003_03101_00001_nrca3_uncal.fits
```

where the first argument is the name of the pipeline configuration file and the second is the name of the input data file.

You can also run the pipeline from within Python by loading and executing the Detector1Pipeline class from the jwst.pipeline package:

```
> from jwst.pipeline import Detector1Pipeline

> result = Detector1Pipeline.call('jw96090001003_03101_00001_nrca3_uncal.fits',
config_file='calwebb_detector1.cfg')
```

In this case the countrate image product ("result") is passed back to the Python shell, so that you can do anything you want with it at that point, including saving it to a file or passing it along as input to subsequent processing. One convenience of processing from within Python is the ability to apply processing to many exposures at once, such as:

```
> from glob import glob

> raw_files = glob('jw96090*_uncal.fits')

> for file in raw_files:

...    result = Detector1Pipeline.call(file, config_file='calwebb_detector1.cfg')

...    result.save(file[:file.rfind('uncal')] + 'rate.fits')
```

CALWEBB_IMAGE2 and CALWEBB_SPEC2

The second stage of calibration processing is applied using the calwebb_image2 or calwebb_spec2 pipelines. calwebb_image2 is designed for use with data from any imaging-like observation mode, such as direct imaging, coronagraphy, Aperture Masking Interferometry (AMI), and all time-series variants of imaging. calwebb_spec2 is to be applied to all spectroscopic modes, including MIRI LRS and MRS, NIRCam and NIRISS Wide-Field Slitless Spectroscopy (WFSS), NIRISS Single Object Slitless Spectroscopy (SOSS), and NIRSpec fixed-slit, MOS, and IFU modes. These two pipeline modules use the _rate.fits or _rateints.fits product from calwebb_detector1 as input, and apply various calibrations to the countrate images, including populating all World Coordinate System (WCS) information, flat-fielding, and attaching or applying flux calibration information. The stage 2 pipelines produce calibrated _cal.fits or _calints.fits products as their output. Complete details on the steps in these pipelines are available in the pipeline reference documents.

These pipelines can be run on individual countrate exposure products (e.g. _rate.fits files) or on an Association of _rate.fits (or rateints.fits) products. An association of countrate products will be used as input, for example, when dedicated off-target background exposures have been obtained as part of the observation and the background exposures are to be subtracted during calwebb_image2 or calwebb_spec2 processing. An ASN file is necessary as input in this case in order to specify the list of multiple exposures (both target and background) and also identify which exposures are to be used as background.

Running calwebb_image2 or calwebb_spec2 on a single exposure can be done from the command-line like this:

```
$ strun calwebb_image2.cfg jw96090001003_03101_00001_nrca3_rate.fits
```

or from within python like this:

```
> from jwst.pipeline import Spec2Pipeline
```

```
> result = Spec2Pipeline.call('jw96090001003_03101_00001_nrca3_rate.fits',
config_file='calwebb_spec2.cfg')
```

Processing an entire association of exposures through these pipelines, using an association (ASN) file as input, can be done like this:

```
$ strun calwebb_image2.cfg jw96090-a3001_20170327t121212_image2_001_asn.json
```

or

```
> Spec2Pipeline.call('jw96090-a3001_20170327t121212_image2_001_asn.json',
config_file='calwebb_spec2.cfg')
```

The ASN file lists one or more output products, each of which is defined to contain one input product (member), hence each input exposure will generate an individual output (calibrated) product. More details on Associations are available in the calibration pipeline package documents.

Stage 3 pipelines

As shown above, there are 5 different stage 3 pipelines, tailored for different observing modes. All stage 3 pipelines must be executed using an Association (ASN) file as input, listing the calibrated exposures to be combined into final products. Details on each of these pipelines are available in the pipeline section of the calibration pipeline package documentation. To run the calwebb_image3 pipeline, for example, you can do:

```
$ strun calwebb_image3.cfg jw96090-a3001_20170327t121212_image3_001_asn.json
```

or from within python:

```
> from jwst.pipeline import Image3Pipeline
```

```
> Image3Pipeline.call('jw96090-a3001_20170327t121212_image2_001_asn.json',
config_file='calwebb_image3.cfg')
```

Pipeline User's Guide

This article contains links to information that will help users understand how to run the JWST calibration pipeline modules and individual calibration steps on their own, should it be desirable to reprocess their data sets.

A detailed User's Guide for the JWST calibration pipeline is available as part of the built-in documents in the jwst software package.

Software Reference Documentation

This article contains links to detailed software reference information for all pipeline modules and individual steps in the JWST calibration pipeline.

Detailed reference documentation for every calibration pipeline module and calibration step is available as part of the built-in documentation in the JWST software package.

Algorithm Documentation

The structure and step-by-step algorithms for the JWST calibration pipeline are documented along with the process that is used for deciding the pipeline algorithms as well as any updates.

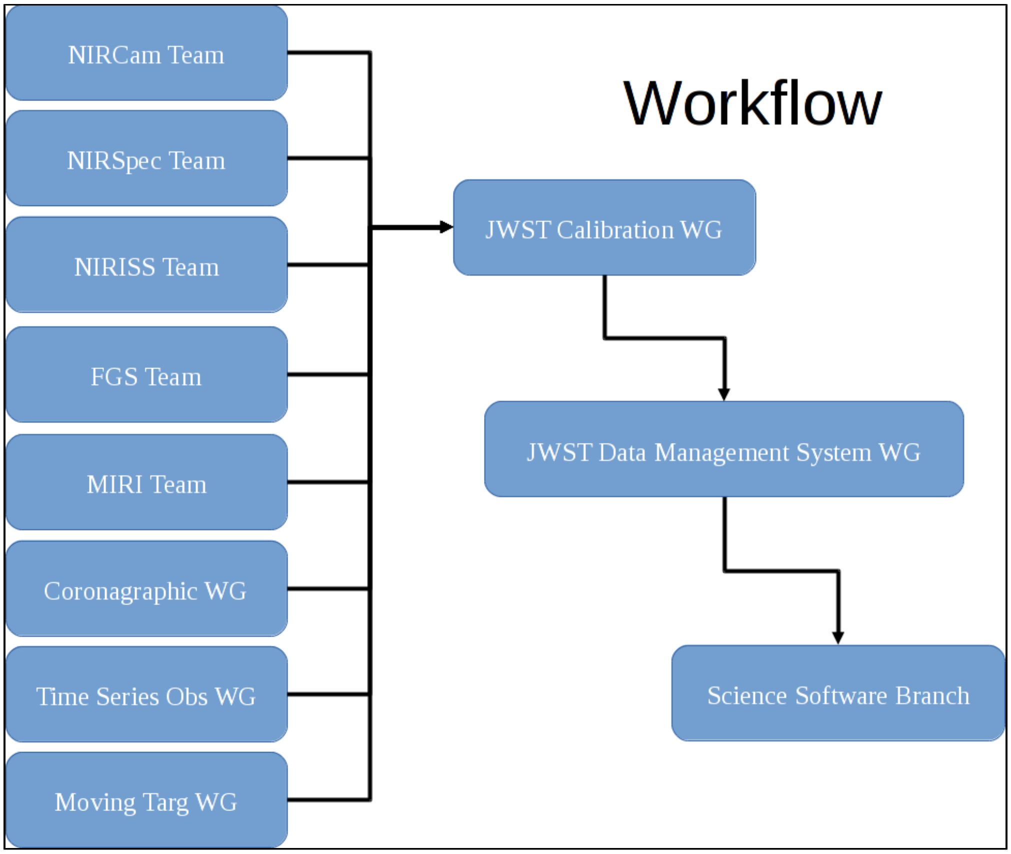
Pipeline organization

The calibration pipeline consists of three stages. Each stage consists of many steps. Where practical, calibration pipeline stages and steps are shared across instruments and observing modes. For example, the near-infrared instruments use similar detectors, allowing similar calibration steps in the initial detector processing stage. In later stages of the calibration pipeline, significantly different steps are needed leading to an increasing divergence in the steps used to process the data (e.g., imaging versus spectroscopy).

Algorithm decisions

The algorithms for use in the calibration pipeline are the result of input from the instrument teams and a few mode specific working groups. These inputs are reviewed and discussed in the JWST Calibration Working Group (CalWG) to ensure that the best possible algorithms. The overall goal of the CalWG is to adopt the best possible algorithms based on experience from previous missions and observatories (Hubble, Spitzer, ground-based) and experience from the ground-testing of the JWST instruments. The membership of the JWST CalWG is composed of all members of the instrument and community that are interested in contributing to the effort of defining the best possible algorithms for JWST. After acceptance by the JWST Cal WG, they are implemented in code by the Science Software Branch with the JWST Data Management Systems WG (DMSWG) ensuring that the requested algorithms are implemented through extensive testing of the as-built pipeline.

Figure 1. JWST calibration pipeline structure



The nominal flow of algorithms is shown. Input is provided by the JWST Instrument Teams and three mode specific working groups (Coronagraphy, Time Series Observations, and Moving Targets). The inputs are reviewed by the JWST CalWG. The resulting algorithms are passed to the JWST DMSWG who ensure that the calibration pipeline coded by the Science Software Branch is as requested.

Baseline version

The goal of the baseline version of the calibration pipeline is to produce data reduction for all instruments and observing modes prior to launch. This focused the efforts of the JWST CalWG to find algorithms for each step in the pipeline that will produce good quality science at launch while at the same time requiring reasonable computational resources.

Optimal version

The optimal version of the calibration pipeline will produce the best quality data reductions. The optimal algorithms for each step start with the baseline and are updated if and when there is a better algorithm proposed. As a result, the optimal version of the pipeline is a never ending goal only approached asymptotically. In other words, the "optimal" name describes the continuing development of the pipeline through the JWST mission.

Algorithms

The details of the algorithms are given by Stages of Processing, Modes of Observing, and Observing Template.

Stages of Processing

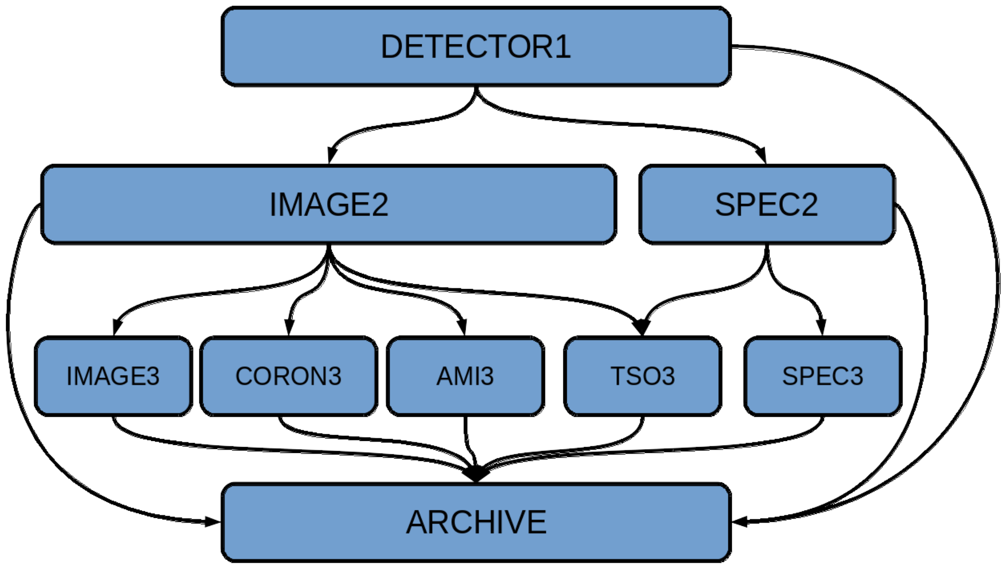
The processing of JWST data goes through three stages. Stage 1 processes the ramps and produces uncalibrated slopes. Stage 2 calibrates the slope images. Stage 3 does processing of ensembles of slopes images.

Pipeline stages

The pipeline has three main stages that provide data to the archive (see Figure 1). Each stage is composed of multiple steps (aka modules). The input and output data are given.

- Stage 1: raw non-destructively read ramps to uncalibrated slope images
  - input: Raw ramps for all integrations.
  - output: Uncalibrated slope images for all integrations and exposures.
  - CALWEBB_DETECTOR1: All data
- Stage 2: calibrates the individual slope images
  - input: Uncalibrated slope images for all integrations and exposures
  - output: Calibrated slope images for all integrations and exposures
  - CALWEBB_IMAGE2: All imaging data (including coronagraphy and AMI)
  - CALWEBB_SPEC2: All spectroscopic data
- Stage 3: ensemble processing of slope images
  - input: Calibrated slope images for all integrations and exposures
  - output: level 3 data
  - CALWEBB_IMAGE3: Direct imaging data
    - output: Coadded image(s) (e.g., mosaic). Catalog of sources. Updated exposure level products.
  - CALWEBB_CORON3: Coronagraphic data
    - output: Coadded image(s). Updated exposure level products.
  - CALWEBB_AMI3: NIRISS Aperture Masking Interferometry data
    - output: Fringe parameters.
  - CALWEBB_SPEC3: Spectroscopic data
    - output: Coadded spectral cube. Extracted 1D spectrum. Updated exposure level products.
  - CALWEBB_TSO3: All time series observations (photometry and spectroscopy)
    - output: Photometry and 1D spectroscopy for each integration and exposure. Updated exposure level products.

Figure 1. JWST pipeline structure



The overall JWST pipeline structure and flow is shown. All the data start at the top and the processing becomes more specific to the particular type of observations as the data flow through the pipeline. All modules in the pipeline contribute processed data to the archive. Note that the "CALWEBB_" prefix has been omitted from the module names.

Stage 1

The 1st stage of the JWST calibration pipeline processes the data from raw ramps to ramp slopes, correcting for instrumental signatures in the process.

The 1st stage of the calibration pipeline takes as input the raw ramps, corrects for a number of detector effects, and outputs the slopes of the ramps.

Stage 1 module:

- CALWEBB_DETECTOR1: All data

CALWEBB_DETECTOR1

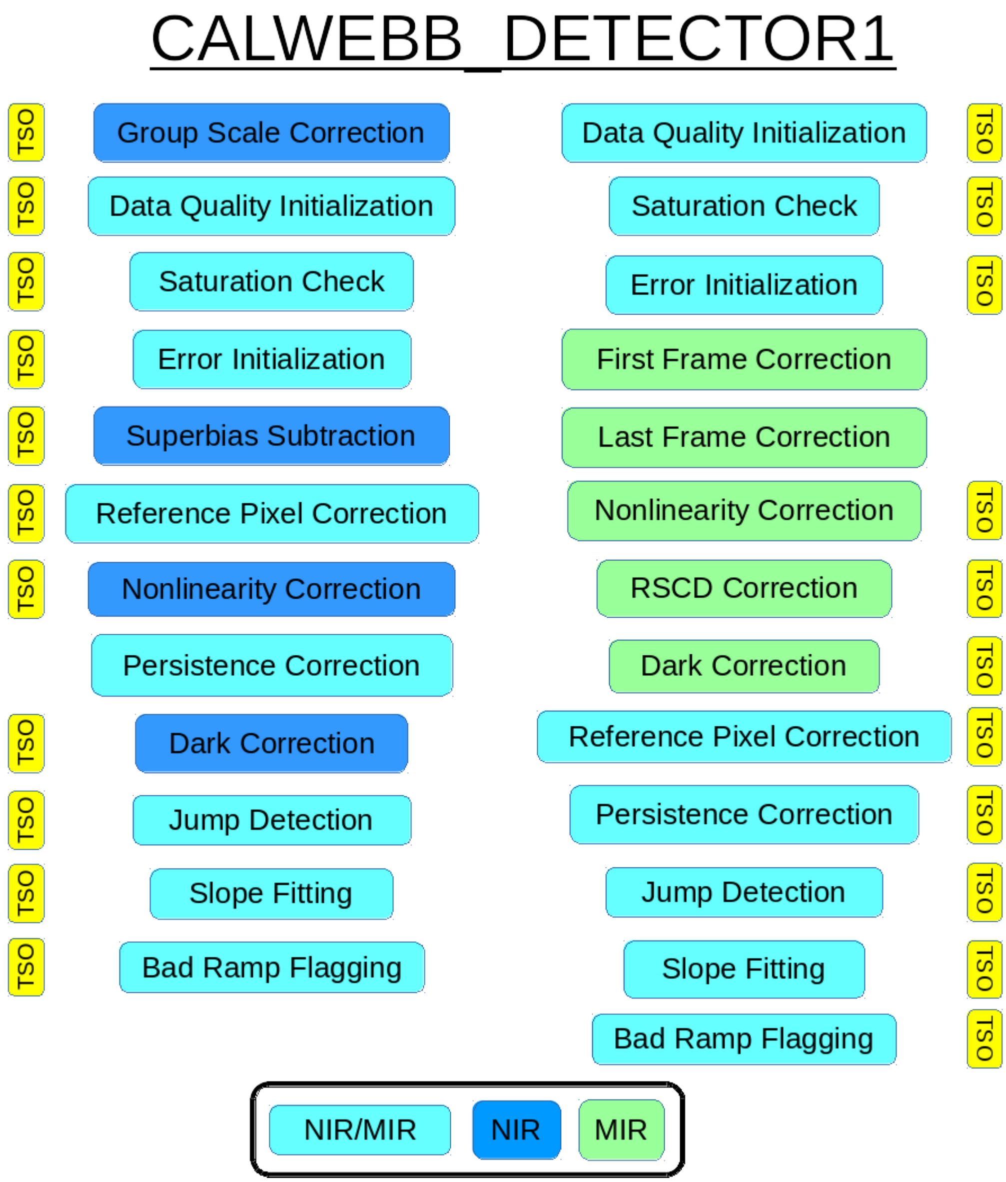
The algorithms for CALWEBB_DETECTOR1, the 1st stage of the JWST calibration pipeline, are described. These algorithms process the data from raw ramps to uncalibrated slopes.

Introduction

The CALWEBB_DETECTOR1 module is the 1st stage in the calibration pipeline and processes data taken with all instruments and modes. The input to this stage is the raw non-destructively read ramps and the output is uncalibrated slope images. A number of instrument signatures are accounted for in this stage and slopes are fit to the corrected ramps. The steps are listed in Figure 1 with the flow from the top to the bottom.

Unless otherwise stated, the algorithms described are the baseline version.

Figure 1. CALWEBB_DETECTOR1



The steps in the CALWEBB_DETECTOR1 module are shown. Those on the left are for the NIR detectors with the right column for the MIR detectors. The common steps are shown in cyan with the NIR and MIR specific steps shown in blue and green respectively. The steps used for Time Series Observations (TSO) are labeled on the far left and right.

NIR/MIR common steps

Data quality initialization

The data quality flags on the pixels are initialized. The pixel data quality flags are initialized based on detector specific reference files and designate permanent conditions that affect all groups for that pixel. Examples of such flags are dead pixels, hot pixels (showing excess dark signal), and low quantum efficiency pixels. The group data quality flags are initialized to zero and designated conditions that only affect some groups for a pixel. Both the pixel and group flags are updated during subsequent pipeline steps as needed.

Saturation check

The analog-to-digital (ADU) values for each group are checked for saturation. The saturation level is set by detector specific reference files. The group data quality flag is set if the DN value in the group is above the reference file set saturation level.

The baseline version of this algorithm does not account for the case where not all of the frames in a group are saturated (some are and some are not). An optimal version of this algorithm that does account for this case is currently under investigation.

Error initialization

The uncertainties are propagated in the pipeline using a noise model. The uncertainty from each step that contributes noise to the final measurement is separately propagated through the pipeline. Different uncertainty sources behave in different ways. Some noise sources (e.g., photon noise) are independent between integrations and others (e.g., flat field noise) are not. In addition, the spatial covariance of different sources varies. By propagating each term through the pipeline, the use of each term can be customized for the processing. For example, the use of the flat field noise term is different between non-dithered and dithered observations. For the former, the noise does not reduce with the addition of more integrations while for the latter it does.

Reference pixel correction

All detectors have reference pixels that have the same readout electronics as regular pixels, but are not sensitive to light. These pixels are located at the edge of the detector arrays and are read out by the same amplifiers as the regular pixels. Thus, reference pixels track drifts in the readout electronics. The average reference pixel level for each readout amplifier is subtracted from the regular pixels for that same amplifier.

Persistence correction

All detectors suffer from persistence giving rise to faint "afterimages" of previous exposures that are seen in the current exposure. The persistence decays exponentially. The persistence is corrected using a trap based model of the persistence with the model details given by detector specific calibration reference files. In addition to correcting the persistence, a pixel flag is set indicating that persistence was corrected at a detectable level.

Jump detection

All the detectors show ramp jumps where the ADU level between two consecutive groups is large relative to those between other consecutive pairs of groups. These ramp jumps are most often caused by cosmic rays that instantaneously deposit large amounts of charge in a pixel. These ramp jumps can be detected and flagged. The number of sigmas above the noise threshold is given as a parameter.

For the baseline algorithm, the two-point difference method is used as it is computationally fast and sufficient for measurements in the photon dominated regime (see Anderson & Gordon 2011). An optimal version of this algorithm that detects smaller ramp jumps in the read noise regime is under investigation.

Slope fitting

The slope for each ramp is determined by performing a weighted linear least squares fit. If a ramp has flagged ramp jumps, the fit is done for each jump free segment and the resulting slopes averaged to produce a single slope per ramp. The weighting is done using the Fixsen et al. (2000) "optimal weighting" method.

An optimal version of this algorithm that uses generalized least squares and a covariance matrix is under investigation.

Bad ramp flagging

There will be ramps that are significantly nonlinear even after all the corrections in CALDETECTOR1 have been done. Ramps that are significantly nonlinear are flagged. The threshold above which to flag a ramp as nonlinear is given in a calibration reference file.

NIR specific steps

Group scale correction

In rare cases, groups can be taken where the the number of frames in a group is not a power of 2. This results in the on-board software not dividing by the appropriate number of frames as the on-board software does averaging by simple bit-shifting. The information on the number of actual frames in a a group and the on-board assumed number of frames in the group is known. This step multiplies all the raw DN values by the appropriate ratio to provide the correct DN values for the actual number of frames averaged per group.

Superbias subtraction

The overall DN level in each pixel and group is offset by a bias level. The nonlinearity in a ramp is relative to this bias level. Thus, the bias level in each group for each pixel is subtracted based on detector specific calibration reference files to provide the correct DN level for the nonlinearity correction step.

Nonlinearity correction

The NIR detectors show a nonlinearity that is due to the gain changing. This non-linearity is well fit with a low order polynomial fit versus relative DN where relative is referenced to the bias level. The correction is done using the fitted polynomials whose parameters are provided by detector specific reference files.

For the baseline algorithm, grouped data is treated the same as non-grouped data and the error due to this assumption is small. An optimal version of this algorithm that accounts for the effects of grouping on the nonlinearity correction is under investigation.

Dark correction

The NIR detectors show excess signal in dark exposures. This excess signal is subtracted group-by-group using detector specific reference files.

MIR specific steps

First frame correction

The 1st frame of MIR data has a transient that has not been fully characterized.

For the baseline algorithm, this 1st frame is flagged and not used further in the pipeline. An optimal version of this algorithm that corrects for this transient is under investigation.

Last frame correction

The MIR detectors show a transient in the last frame that is caused by the reset pattern. The last frame of the MIR detectors is a read-reset frame where two rows are read and then reset. Then the next two rows are processed in the same manner. The transient is caused by reset strongly changing the values in the to-be-read pixels due to the coupling between pixels.

For the baseline algorithm, the last frame is flagged and not used further in the pipeline.

An optimal version of this algorithm is under development. The amplitude of the transient seems to be well characterized by a polynomial fit to the value in the last frame itself. The correction under development will use polynomial parameters provided in detector specific reference files.

Nonlinearity correction

The MIR detectors show a non-linearity that is due to the quantum efficiency changing. This non-linearity is well fit with a low order polynomial fit versus absolute DN. The non-linearity is wavelength dependent at wavelengths above approximately 21 microns. The correction is done using the fitted polynomials whose parameters are provided by detector and wavelength specific reference files.

RSCD correction

The Reset Switch Charge Decay is a transient seen in the 2nd and subsequent integrations in a MIR exposure. If uncorrected, this transient results in the slopes in the 2nd+ integrations to be larger than the 1st integration. This transient is well described by a decaying single exponential that is proportional to the counts at the end of the previous integration. The correction involves subtracting this exponential from the 2nd+ integration ramps where the parameters of the exponential are set by detector specific reference files.

Dark correction

The MIR darks show a dependence on the integration number in an exposure. The dark subtraction is done group-by-group using detector and integration specific reference files.

References

Anderson & Gordon 2011, PASP, 123, 1237
Optimal Cosmic-Ray Detection for Nondestructive Read Ramps

Fixsen et al. 2000, PASP, 112, 1350
Cosmic-Ray Rejection and Readout Efficiency for Large-Area Arrays

JWST technical documents

Stage 2

The 2nd stage of the JWST calibration pipeline processes the uncalibrated slope images to calibrated slope images correcting for some detector/instrument/telescope effects in the process.
The 2nd stage of the calibration pipeline takes as input uncalibrated slope images, corrects for some detector/instrument/telescope effects, and outputs calibrated slope images.

Stage 2 modules:

- CALWEBB_IMAGE2: All imaging data (including coronagraphy and AMI)
- CALWEBB_SPEC2: All spectroscopic data

CALWEBB_SPEC2

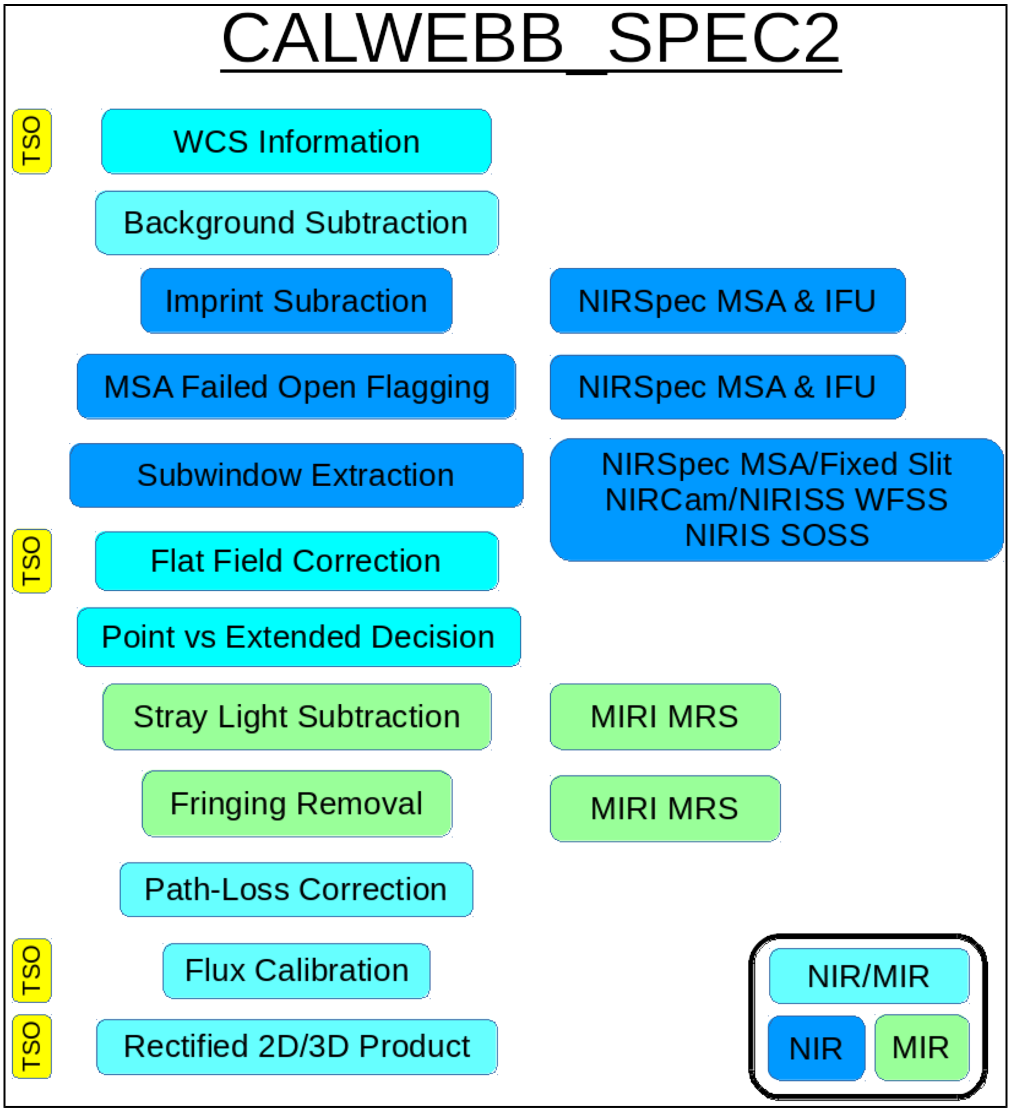
The algorithms for CALWEBB_SPEC2, the 2nd stage of the JWST calibration pipeline for spectroscopic data, are described. These algorithms process the data from uncalibrated slope images to calibrated slope images.

Introduction

The CALWEBB_SPEC2 module is the 2nd stage in the calibration pipeline for all spectroscopic observations. The input to this stage is the uncalibrated slope images (CALWEBB_DETECTOR1 output) and the output is calibrated slope images.The steps are listed in Figure 1 with the flow from the top to the bottom.

Unless otherwise stated, the algorithms described are the baseline version.

Figure 1. CALWEBB_SPEC2



The steps in the CALWEBB_SPEC2 module are shown. The common steps are shown in cyan with the NIR and MIR specific steps shown in blue and green respectively. The steps used for Time Series Observations (TSO) are labeled on the far left. The specific instrument modes are listed for some of the steps in the right column.

NIR/MIR steps

WCS information

The information to transfer the pixel coordinates to astronomical coordinates (e.g., ra, dec, and wavelength) is added to the data with this step. The needed information is described in the World Coordinate System (WCS) format. The WCS information and distortion model are provided by instrument and detector specific calibration reference files (all) and an instrument model (NIRSpec only). The data itself is not modified by this step.

Background subtraction

An observed or modeled background image is subtracted from the target exposure.

If an APT specified background target was observed, a combined background image constructed from all exposures of that target is subtracted from the science target exposure.

For NIRISS WFSS, NIRISS SOSS, and NIRCam WFSS, a master version of the dispersed background is used, scaled to match the background in the science target exposure.

For NIRSpec MSA, NIRSpec IFU, NIRSpec Fixed Slit, and MIRI LRS, if observations were taken at two or more nod positions, the associated nod position or sum of nod positions is subtracted.

Flat field correction

The flat field corrects for the pixel-to-pixel variations and large scale variations in the instrument+telescope responsivity. The flat field image is taken from instrument and detector specific calibration reference files. For NIRSpec, the flat field is created on-the-fly from reference files and the instrument model.

Point vs extended decision

Some of the pipeline steps rely on knowing if the target is a point or extended source (e.g., spectral extraction). The determination of a source as point or extended can be informed by setting the extended tag in APT (see JWST APT Targets), from simultaneous imaging (NIRISS and NIRCam WFSS), from the NIRSpec MSA tool output, or if no other information is available, by using defaults by instrument mode. The default is to assume a point source except for NIRSpec MSA backgrounds and MIRI IFU observations. This decision is attached to the data and used by later steps in the pipeline. When manually rerunning the calibration pipeline, this step can be set by the user.

Path-loss correction

The loss of signal in the path is corrected here. The causes of signal loss include the slit (NIRSpec, MIRI), the diffracted slit image being larger than the gratings (NIRSpec), the MSA bar shadow (NIRSpec MSA), and subarrays smaller than diffracted slit image (NIRISS SOSS). The correction values are provided by instrument and mode specific calibration reference files

Flux calibration

The multiplicative conversion factor between counts/sec and MJy/sr as a function of wavelength is attached to the data. A second conversion factor between counts/sec and micro-Jy/sq arc sec as a function of wavelength is also attached. Basically, these are FITS keywords. The pixel area reference file will be attached to the data allowing conversion between surface brightness and flux density for each pixel. Details of the calculation of these calibration factors can be found at Absolute Flux Calibration.

Rectified 2D/3D product

As a product for the archive, rectified 2D (all except IFUs) or 3D (IFUs) products are created using the attached WCS information. These rectified 2D/3D product are not used in the pipeline itself, and are created as they are useful for visual browsing of the data.

NIR steps

Imprint subtraction

The MSA shutters are not completely dark allowing a small amount of light to leak through causing an imprint. When there is a dedicated imprint exposure taken (all MSA shutters closed and IFU closed), this will be subtracted from the target exposure.

MSA failed open flagging

The MSA failed open shutters will result in an elevated level of light to fall on the detectors. This step flags the pixels that are affected by these failed open shutters. This step uses the list of failed open shutters and the NIRSpec instrument model to flag the affected pixels.

Subwindow extraction

The region of interest for each source is extracted for NIRISS/NIRCam WFSS, NIRISS SOSS, and NIRSpec MSA/Fixed Slit data. The location of the sources for the WFSS observations is provided using the direct imaging observations taken as part of the observations. The NIRSpec MSA configuration file provides the location of the sources for this MOS mode. The other modes have fixed locations for sources. This subwindow extraction assumes that sources are isolated and not confused with other sources (such confused sources will need to be extracted with a post-pipeline tool). The main impact of this step is to provide single sources to the rest of the pipeline providing a uniform treatment of all spectra from the observation modes listed.

MIR steps

Stray light subtraction

There is stray light present in between the slices in the MIRI IFU observations. This stray light is subtracted by linearly interpolating the measured surface brightnesses between the inter-slice regions.

Fringing removal

There are significant fringes in the MIRI IFU data. These fringes are removed (to first order) by dividing the exposures by the detector specific calibration reference file.

CALWEBB_IMAGE2

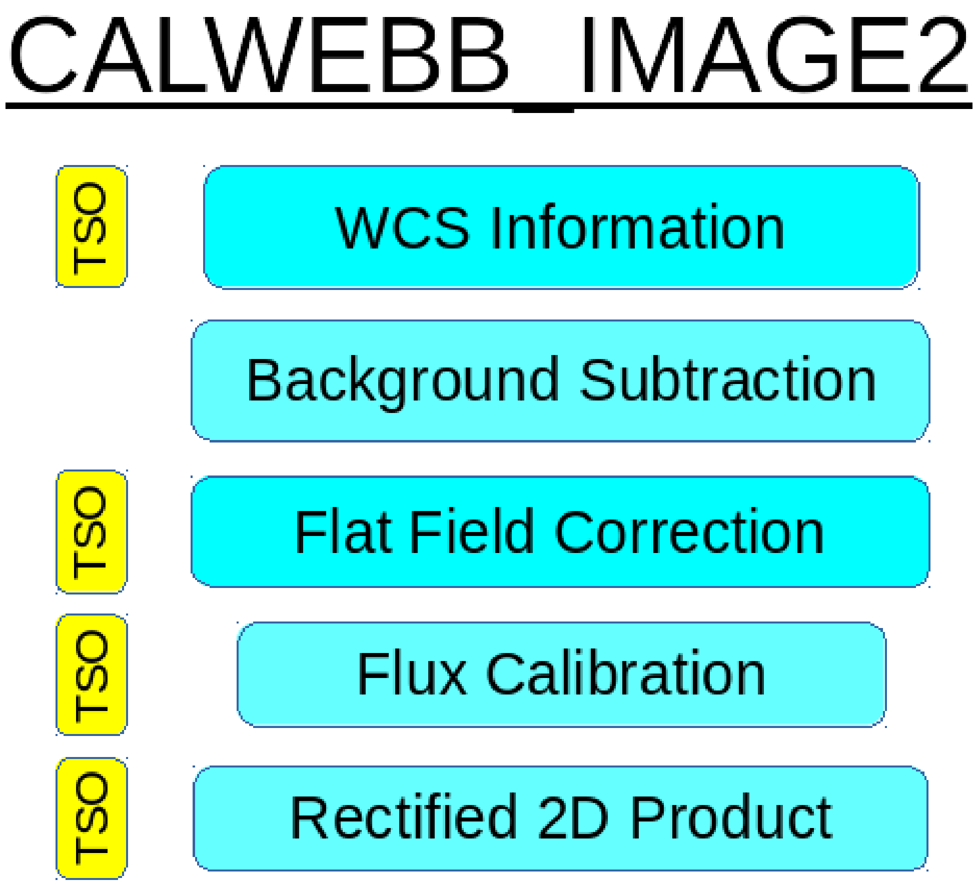
The algorithms for CALWEBB_IMAGE2, the 2nd stage of the JWST calibration pipeline for imaging data, are described. These algorithms process the data from uncalibrated slope images to calibrated slope images.

Introduction

The CALWEBB_IMAGE2 module is the 2nd stage in the calibration pipeline for all imaging observations. The input to this stage is the uncalibrated slope images (CALWEBB_DETECTOR1 output) and the output is calibrated slope images. The steps are listed in Figure 1 with the flow from the top to the bottom.

Unless otherwise stated, the algorithms described are the baseline version.

Figure 1. CALWEBB_IMAGE2



The steps in the CALWEBB_IMAGE2 module are shown.

NIR/MIR steps

There are only common NIR/MIR steps for this module.

WCS information

The information to transfer the pixel coordinates to astronomical coordinates (e.g., ra & dec) is added to the data with this step. The needed information is described in the standard World Coordinate System (WCS) format—basically standard FITS WCS keywords and a distortion model. The WCS information and distortion model are provided by instrument and detector specific calibration reference files. The image data itself is not modified by this step.

Background subtraction

If an APT specified background target was observed, a combined background image constructed from all exposures of that target is subtracted from the science target exposure. The background observations can be specified by the user when the calibration pipeline is being manually rerun.

Flat field correction

The flat field corrects for the pixel-to-pixel variations and large scale variations in the instrument+telescope responsivity. The flat field image is taken from instrument and detector specific calibration reference files.

Flux calibration

The multiplicative conversion factor between counts/sec and MJy/sr is attached to the data. A second conversion factor between counts/sec and micro-Jy/sq arcsec is also be attached. Basically, these are FITS keywords. The pixel area reference file is also attached to the data allowing conversion between surface brightness and flux density for each pixel. Details of the calculation of these calibration factors can be found at Absolute Flux Calibration.

Rectified 2D product

As a product for the archive, rectified 2D images are created using the AstroDrizzle algorithm and the attached WCS information. This rectified 2D product is not used in the pipeline itself, it is useful for visual browsing of the data.

Stage 3

The 3rd stage of the JWST calibration pipeline processes the calibrated slope images to produce combined products (e.g., mosaics and spectral cubes) and extracted data (e.g., source photometry and extracted spectra).

The 3rd stage of the calibration pipeline takes as input calibrated slope images, performs processing using the ensemble of images, and produces combined products (e.g., mosaics and spectral cubes) and extracted data (e.g., source photometry and extracted spectra).

Stage 3 modules are:

- CALWEBB_IMAGE3: Direct imaging data
- CALWEBB_CORON3: Coronagraphic data
- CALWEBB_AMI3: NIRISS Aperture Masking Interferometry data
- CALWEBB_SPEC3: Spectroscopic data
- CALWEBB_TSO3: All time series observations (photometry and spectroscopy)

CALWEBB_AMI3

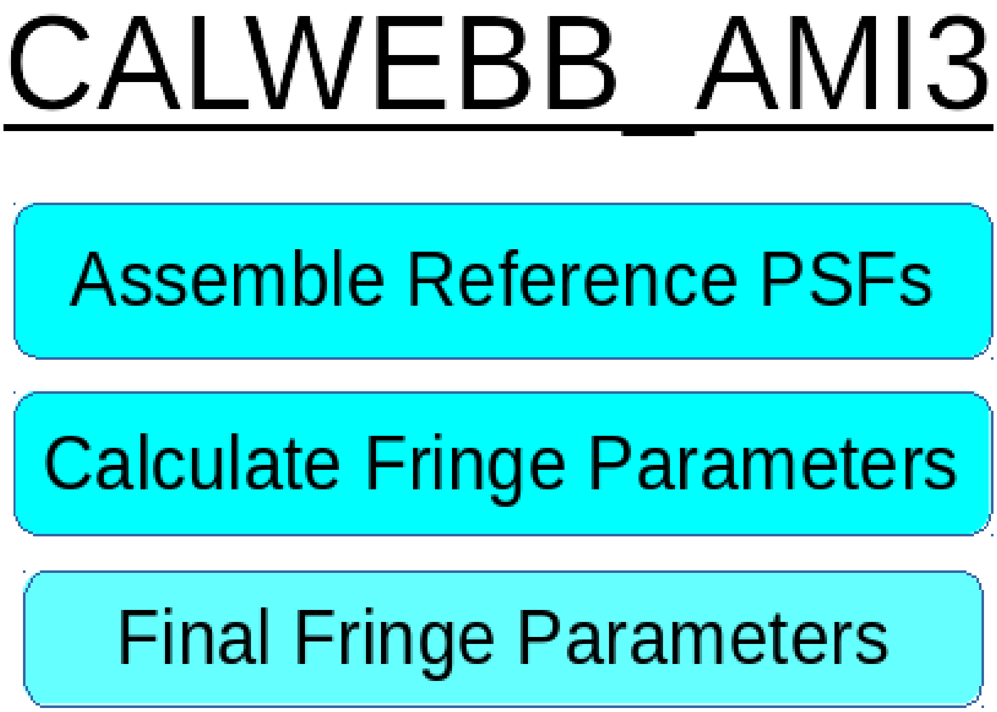
The algorithms for CALWEBB_AMI3, the 3rd stage of the JWST calibration pipeline for NIRISS Aperture Masking Interferometry (AMI), are described. These algorithms process the data from calibrated slope images to fringe parameters.

Introduction

The CALAMI3 module is the 3rd stage in the calibration pipeline for the NIRISS Aperture Masking Interferometry (AMI) data. The inputs to this stage are the calibrated slope images (CALWEBB_IMAGE2 output) and the output is fringe parameters. The steps are listed in Figure 1 with the flow from the top to the bottom.

Unless otherwise stated, the algorithms described are the baseline version.

Figure 1. CALWEBB_AMI3



The steps in the CALWEBB_AMI3 module are shown.

Steps

Assemble reference PSFs

The available list of reference PSFs is generated. These reference PSFs consist of reference stars specifically observed for the target.

Calculate fringe parameters

The closure phases and visibilities for each reference PSF and target are calculated. As part of this calculation, the centers of the reference and target PSFs are determined (iteratively).

Final fringe parameters

The reference PSF closure phases are subtracted from the target closure phases. The target visibilities are divided by the reference PSF visibilities. These two calculations are done matched by dither position.

CALWEBB_CORON3

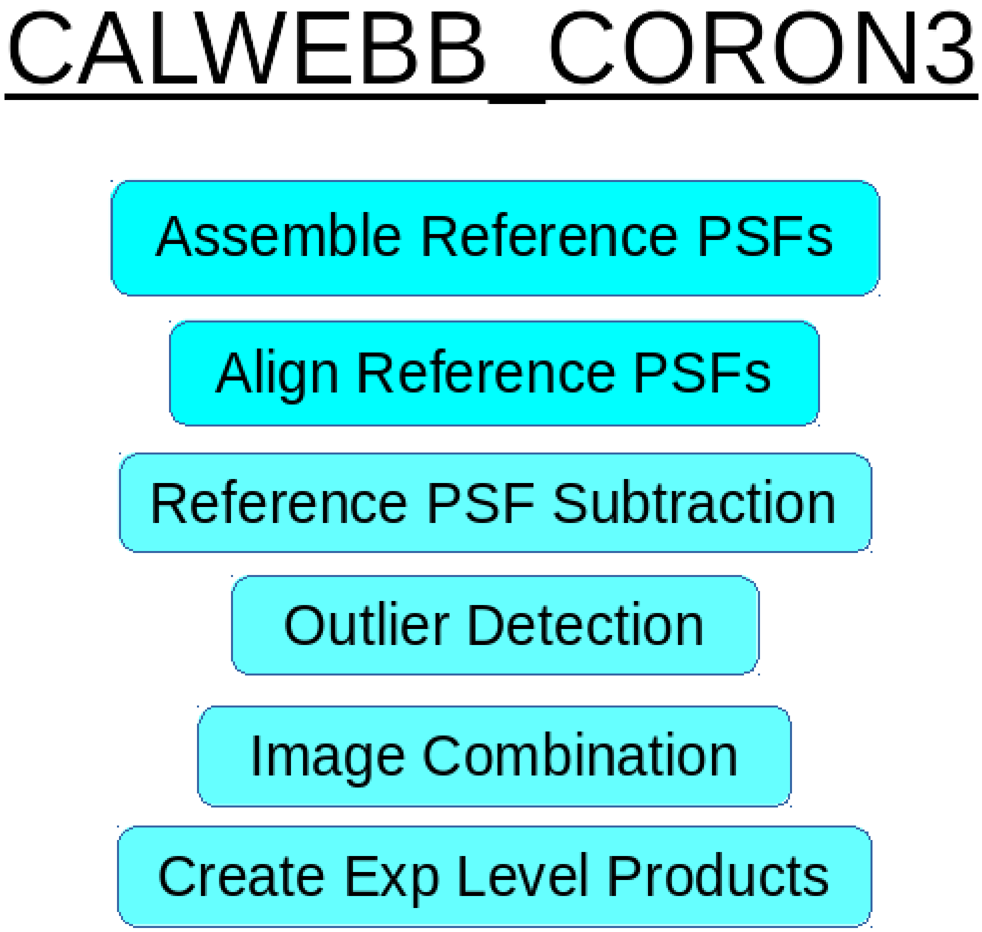
The algorithms for CALWEBB_CORON3, the 3rd stage of the JWST calibration pipeline for coronagraphic data, are described. These algorithms process the data from calibrated slope images to combined, reference PSF subtracted coadded images.

Introduction

The CALWEBB_CORON3 module is the 3rd stage in the calibration pipeline for the NIRCam and MIRI Coronagraphic data. The inputs to this stage are the calibrated slope images (CALWEBB_IMAGE2 output) and the output is a reference PSF subtracted image stack. The steps are listed in Figure 1 with the flow from the top to the bottom.

Unless otherwise stated, the algorithms described are the baseline version.

Figure 1. CALWEBB_CORON3



The steps in the CALWEBB_CORON3 module are shown.

NIR/MIR steps

Assemble reference PSFs

The available list of reference PSFs is generated. These reference PSFs can include reference stars specifically observed for the target as well as observations of the target taken at different orientations.

Align reference PSFs

The reference PSFs are aligned with the target observation using the Fourier LSQ algorithm to measure the shifts and the Fourier Shift algorithm to apply the shifts to each reference PSF integration.

Reference PSF subtraction

The reference PSF that is subtracted from each target integration is created using the list of reference PSFs and the KLIP algorithm (Soummer et al. 2012).

Outlier detection

A 2nd pass at outlier detection is done using the image stack. The majority of the outliers will be due to cosmic rays undetected during the 1st pass at outlier detection done in CALWEBB_DETECTOR1. An iterative sigma clipping algorithm is used in pixel coordinates on the image stack. The presence of an outlier results in a pixel flag being set.

Image combination

The target images (including those at different rotations) are combined into a single combined image. This is done using the AstroDrizzle code with the output pixel size set to the input pixel size.

Create exposure level products

The exposure level products are re-created at this stage to provide the highest quality products that include the results of the ensemble processing (updated WCS, matching backgrounds, and 2nd pass outlier detection). These products are for the archive and include the unrectified 2D images and rectified 2D images.

References

Soummer, R., Pueyo, R., Larkin, J., 2012, ApJ, 755, 28
Detection and Characterization of Exoplanets and Disks Using Projections on Karhunen-Loève Eigenimages

JWST technical documents

CALWEBB_IMAGE3

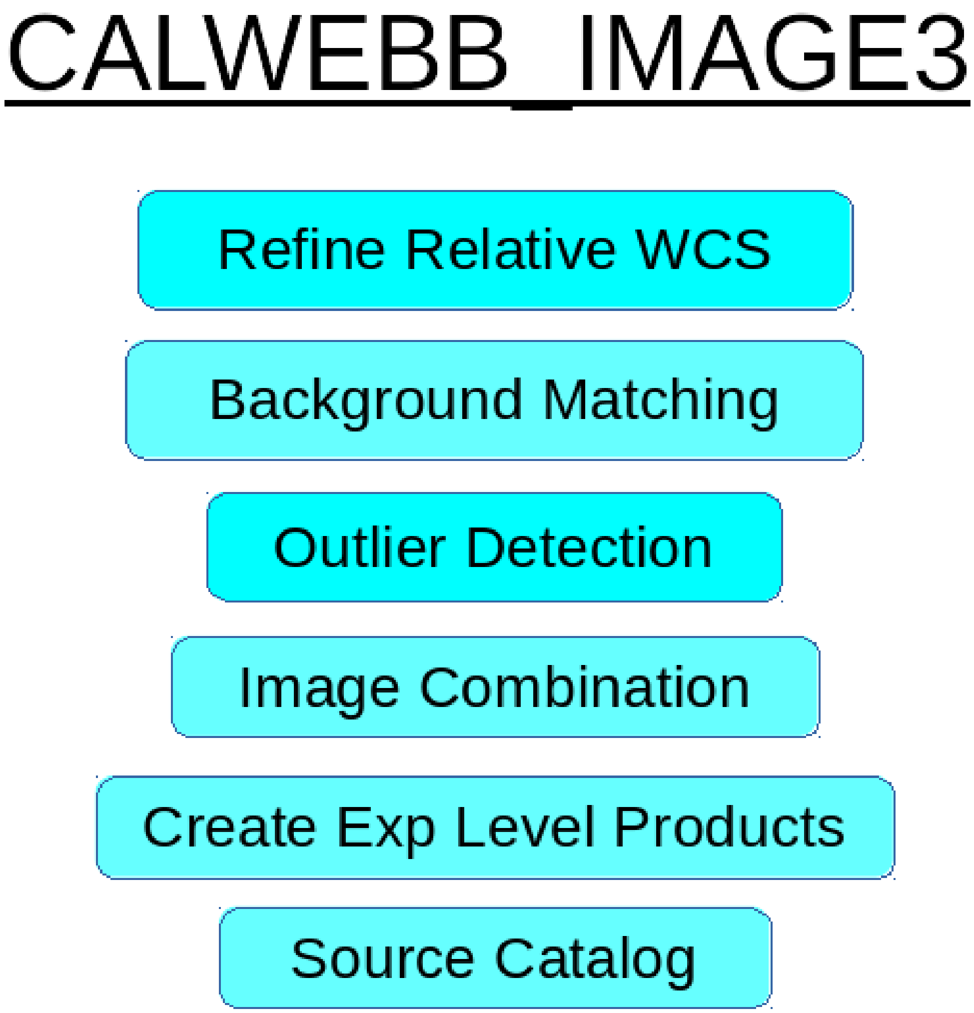
The algorithms for CALWEBB_IMAGE3, the 3rd stage of the JWST calibration pipeline for imaging data, are described. These algorithms process the data from calibrated slope images to mosaics and source catalogs.

Introduction

The CALWEBB_IMAGE3 module is the 3rd stage in the calibration pipeline for all imaging data. The inputs to this stage are the calibrated slope images (CALWEBB_IMAGE2 output) and the outputs are mosaics and source catalogs. The steps are listed in Figure 1 with the flow from the top to the bottom.

Unless otherwise stated, the algorithms described are the baseline version.

Figure 1. CALWEBB_IMAGE3



The steps in the CALWEBB_IMAGE3 module are shown.

NIR/MIR steps

Refine relative WCS

The relative WCS information between the different images is refined using the locations of common sources between images. The AstroDrizzle code is used for this step. The WCS information associated with each image is updated by this step.

Background matching

The background levels in images can vary as a function of time due to the thermal telescope emission, zodiacal emission, etc. This step corrects the overall background level of each image so that the overlapping regions of the images have the same background. The AstroDrizzle code is used for this step.

Outlier detection

A 2nd pass at outlier detection is done using the overlapping regions observed in different exposures. The majority of the outliers will be due to cosmic rays undetected during the 1st pass at outlier detection done in CALWEBB_DETECTOR1. The AstroDrizzle code is used for this step. The presence of an outlier results in a pixel flag being set.

Imaging combination

The combination of multiple images into a mosaic is done using the AstroDrizzle code with the output pixel size set to the input pixel size. This is done for all the images taken in a specific band in an observation. Pixels flagged as "bad" (outliers, hot pixels, etc.) are not used in the creation of the mosaic.

Note that moving targets are supported and the mosaics are created in the moving target reference frame.

Create expourse level products

The exposure level products are re-created at this stage to provide the highest quality products that include the results of the ensemble processing (updated WCS, matching backgrounds, a 2nd pass outlier detection). These products are for the archive and include the unrectified 2D images and rectified 2D images.

Source catalog

Source catalogs are created using the Astropy Photutils package. The goal of this step is to produce a good quality catalog that will provide a 1st pass at such a catalog. This catalog is needed for the reduction of NIRISS/NIRCam WFSS observations and to setup NIRSpec MSA observations.

CALWEBB_SPEC3

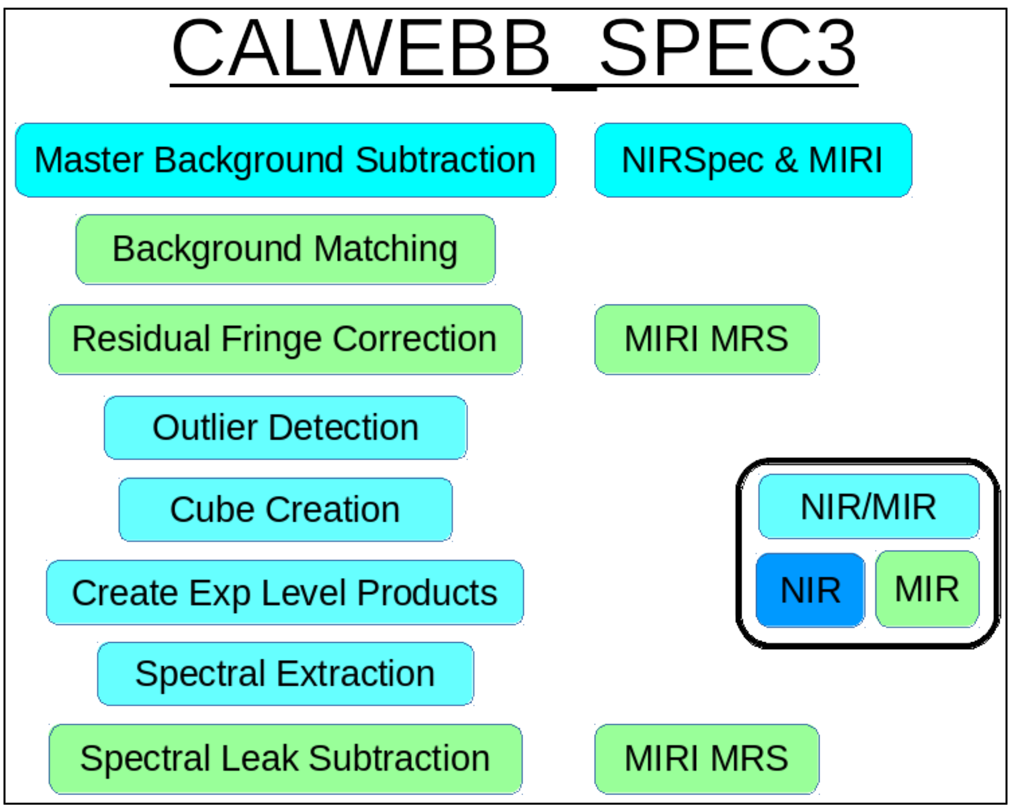
The algorithms for CALWEBB_SPEC3, the 3rd stage of the JWST calibration pipeline for imaging data, are described. These algorithms process the data from calibrated slope images to combined spectral data (2D or 3D) and extracted spectra.

Introduction

The CALWEBB_SPEC3 module is the 3rd stage in the calibration pipeline for spectroscopic data. The inputs to this stage are the calibrated slope images (CALWEBB_SPEC2 output) and the output is combined 2D images/3D spectral cubes and extracted spectra.The steps are listed in Figure 1 with the flow from the top to the bottom.

Unless otherwise stated, the algorithms described are the baseline version.

Figure 1. CALWEBB_SPEC3



The steps in the CALWEBB_SPEC3 module are shown. The dark-lined box on the right is a legend that shows how the instrument(s) are represented by color. Steps that are common to NIR and MIR are shown in cyan. Steps specific to NIR are in blue, while steps for MIR are in green.

NIR/MIR steps

Master background subtraction

The background can be improved over the CALWEBB_SPEC2 step under the assumption that the background does not change across the field of view. With this assumption, a high signal-to-noise spectrum of the background can be measured by averaging over the spatial dimension(s). Then the master background spectrum can be used to create the equivalent 2D background image for subtraction from the target exposures. This step would be done instead of the CALWEBB_SPEC2 background subtraction step. When to do master background subtraction versus background subtraction is still to-be-decided. Hence this step is an optimal pipeline step.

Outlier detection

A 2nd pass at outlier detection is done using the overlapping regions observed in different exposures. The majority of the outliers will be due to cosmic rays undetected during the 1st pass at outlier detection done in CALWEBB_DETECTOR1. The presence of an outlier results in a pixel flag being set.

Cube creation

Combinations of dithered spectral observations are done in 2D and 3D as appropriate. 2D images (spectral versus spatial) are created for NIRSpec Fixed Slit/MIRI LRS nodded observations, NIRSpec MSA data taken with shutters above/below, and NIRISS/NIRCam WFSS dithered observations. 3D images (spectral versus ra/dec) are created for NIRSpec and MIRI IFU and MIRI LRS observations taken with cross-slit steps. The combination is done in a single step/interpolation to avoid propagating noise through multiple interpolations.

Note that moving targets are supported and the cubes are created in the moving target reference frame.

Create exposure level products

The exposure level products are re-created at this stage to provide the highest quality products that include the results of the ensemble processing (updated WCS, matching backgrounds, a 2nd pass outlier detection). These products are for the archive and include the rectified 2D (spatial versus spectral, all except IFUs) or 3D (spectral cubes, IFUs) products

Spectral extraction

For point sources, the extraction is done using a boxcar (2D data) or circular aperture (3D data) with local background subtraction measured from apertures outside the extraction aperture.

For extended sources, the extraction is done with rectangular apertures with no local background subtraction.

The extraction of spectra is done from the rectified observations for the baseline pipeline. An optimal version of spectral extraction that extracts directly from the unrectified data is under investigation.

MIR steps

Background matching

The background levels in observations can vary as a function of time due to the thermal telescope emission, zodiacal emission, etc. This step corrects the overall level of each spectral cube so that the overlapping regions between exposures have the same background. This step is only done for MIRI IFU observations.

Residual fringe correction

The correction for fringes in CALWEBB_SPEC2 uses a static calibration reference file. The fringes change slightly between observations due to the exact placement of the source in the IFU slices. A residual fringe correction is made by fitting the residual fringes in each spatial pixel using a model that is constrained to only have certain spatial frequencies. This step is only done for MIRI IFU observations.

Spectral leak subtraction

The MIRI IFU observations have a spectral leak in channel 3A at 12.2 µm due to 2nd order light at 6.1 µm with a throughput of ~2%. If observations exist for band 1B, an appropriately scaled 1B extracted spectrum is subtracted from the extracted 3A spectrum.

CALWEBB_TSO3

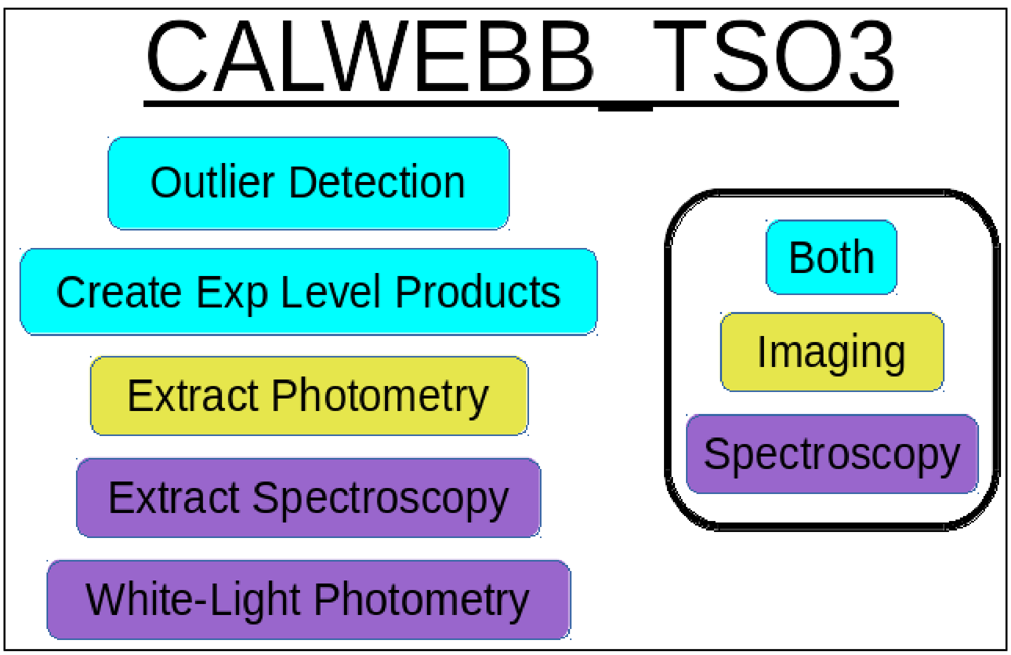
The algorithms for CALWEBB_TSO3, the 3rd stage of the JWST calibration pipeline for all Time Series Observations (TSO) data, are described. These algorithms process the data from calibrated slope images to extracted photometry and spectroscopy for each integration.

Introduction

The CALWEBB_TSO3 module is the 3rd stage in the calibration pipeline for the all Time Series Observation (TSO) data. The inputs to this stage are the calibrated slope images (CALWEBB_IMAGE2 and CALWEBB_SPEC2 output) and the outputs are extracted photometry and spectroscopy as appropriate.The steps are listed in Figure 1 with the flow from the top to the bottom.

Unless otherwise stated, the algorithms described are the baseline version.

Figure 1. CALWEBB_TSO3



The steps in the CALWEBB_TSO3 module are shown. The common steps are shown in cyan with the Imaging and Spectroscopy specific steps shown in yellow and purple respectively.

Common steps

Outlier detection

A 2nd pass at outlier detection is done using the image stack. The majority of the outliers will be due to cosmic rays undetected during the 1st pass at outlier detection done in CALDETECTOR1. An iterative sigma clipping algorithm is used in pixel coordinates on the image stack. The presence of an outlier results in a pixel flag being set.

Create exposure level products

The exposure level products are re-created at this stage to provide the highest quality products that include the results of the ensemble processing (a 2nd pass outlier detection). These products are for the archive and include the unrectified 2D images and rectified 2D images.

Imaging steps

Extract photometry

Aperture photometry with a predefined aperture is performed for each integration providing the time series photometry.

Spectroscopy steps

Extract spectroscopy

The extraction is done using a boxcar for each integration providing time series spectroscopy.

White-light photometry

The 1D extracted spectra are collapsed to create time series white-light photometry for the spectral observations.

Modes of Observing

The calibration pipeline stages run for each type of JWST observation and observing template are given. The calibration pipeline modules are given as a function of the type of observations and Observing Template.

Type	Observing templates	Pipeline modules
Imaging	NIRCam imaging NIRISS imaging MIRI imaging	CALWEBB_DETECTOR1 CALWEBB_IMAGE2 CALWEBB_IMAGE3
Spectroscopy	NIRCam WFSS NIRSpec MOS NIRSpec FS spectroscopy NIRSpec IFU spectroscopy NIRSpec TSO spectroscopy NIRISS WFSS MIRI LRS MIRI MRS	CALWEBB_DETECTOR1 CALWEBB_SPEC2 CALWEBB_SPEC3
Coronagraphy	NIRCam coronagraphic imaging MIRI coronagraphic imaging	CALWEBB_DETECTOR1 CALWEBB_IMAGE2 CALWEBB_CORON3
Aperture Masking Interferometry	NIRISS AMI	CALWEBB_DETECTOR1 CALWEBB_IMAGE2 CALWEBB_AMI3
Time Series Observations: Imaging	NIRCam TSO imaging MIRI imaging (TSO option)	CALWEBB_DETECTOR1 CALWEBB_IMAGE2 CALWEBB_TSO3
Time Series Observations: Spectroscopy	NIRCam TSO grism NIRSpec TSO spectroscopy NIRISS SOSS MIRI LRS (TSO option)	CALWEBB_DETECTOR1 CALWEBB_SPEC2 CALWEBB_TSO3
All Target Acquisition Imaging FGS Science/Guide Imaging		CALWEBB_DETECTOR1 CALWEBB_IMAGE2

By Observing Template

The calibration pipeline modules used to process the data from each JWST Observing Template are given. The relationship between the calibration pipeline modules and Observing Templates is given.

Pipeline Module	Observing Template
CALWEBB_DETECTOR1	All
CALWEBB_IMAGE2	NIRCam Imaging NIRCam Coronagraphic Imaging NIRCam TSO Imaging NIRISS AMI NIRISS Imaging MIRI Imaging MIRI Coronagraphic Imaging All Target Acquisition Imagings FGS Science/Guide Imaging
CALWEBB_SPEC2	NIRCam WFSS NIRCam TSO Grism NIRSpec MOS NIRSpec FS spectroscopy NIRSpec IFU spectroscopy NIRSpec TSO spectroscopy NIRISS SOSS NIRISS WFSS MIRI LRS MIRI MRS
CALWEBB_IMAGE3	NIRCam Imaging NIRISS Imaging MIRI Imaging

CALWEBB_CORON3	NIRCam Coronagraphic Imaging MIRI Coronagraphic Imaging
CALWEBB_AMI3	NIRISS AMI
CALWEBB_SPEC3	NIRCam WFSS NIRSpec MOS NIRSpec FS Spectroscopy NIRSpec IFU Spectroscopy NIRISS WFSS MIRI LRS MIRI MRS
CALWEBB_TSO3	NIRCam TSO Imaging NIRCam TSO Grism NIRSpec TSO Spectroscopy NIRISS SOSS MIRI Imaging (TSO option) MIRI LRS (TSO option)

Obtaining and Installing Software

The JWST calibration pipeline software is available for installation from AstroConda, which is a free Conda channel maintained by the Space Telescope Science Institute (STScI) in Baltimore, Maryland. This channel provides tools and utilities required to process and analyze data from the Hubble Space Telescope (HST), James Webb Space Telescope (JWST), and others.

What's Conda?

Conda is an open source package management system and environment management system for installing multiple versions of software packages and their dependencies and switching easily between them. It works on Linux, OS X and Windows, and was created for Python programs but can package and distribute any software. On-line Conda documentation provides full details on the inner workings of this management system. Briefly, Conda is a package manager application that quickly installs, runs, and updates packages and their dependencies. The conda command is the primary interface for managing installations of various packages. It can query and search the package index and current installation, create new environments, and install and update packages into existing conda environments.

You must have the Conda software installed before you can install the JWST package from the AstroConda channel. Full details on downloading and installing Conda can be found here. In a nutshell, you:

- Download the appropriate version of Anaconda for your system's OS from the Anaconda download page
- Follow the directions on the download page for installing the package on your system

NOTE: Command-line installation and operation of Conda requires use of the bash shell on Linux and OS X systems.

AstroConda

Setup

In order to use Conda to install packages directly from the STScI AstroConda channel you will need to append our URL to Conda’s channel search path.

```
$ conda config --add channels http://ssb.stsci.edu/astroconda
# Writes changes to ~/.condarc
```

The package management system Conda is now configured to pull from our repository, so you may go ahead and install the `stsci` package. This package installs nearly all of the software provided by STScI in one shot, including the JWST calibration pipeline package.

Installation

The following example generates a new conda environment on your system named “astroconda”, however this naming convention is merely a suggestion. Feel free to use a name that works best for you.

```
$ conda create -n astroconda stsci
```

Using the Software

After the installation is complete go ahead and activate the “astroconda” environment. This command only needs to be executed one time per terminal session.

```
$ source activate astroconda
```

Anytime that you terminate and reactivate a terminal window or session and want to use the environment again, you will need to reactivate the environmen, using the above command.

To deactivate the “astroconda” environment or switch to another conda environment on your system, close your terminal window or run:

```
$ source deactivate
```

You can then activate any other environment on your system.

Updates

From time to time STScI will make updates available for the JWST package and dependent packages. You can download and install updates by updating your conda environment with:

```
$ conda update -n astroconda stsci
```

where "astroconda" is the name of the environment on your system and "stsci" is the package to be updated. Conda will automatically search for and install updated versions of all packages contained in "stsci".

JWST File Names, Formats, and Data Structures

JWST data files are FITS files with unique names that map to the original proposal, observation, visit, and instrument and detector used. Their particular format depends on the stage of the JWST Data Reduction Pipeline where these were created.
All JWST data shares a basic data structure with slight variations that depend on the observing mode or instrument used. Being able to work with JWST data requires an understanding of JWST FITS files, Advanced Scientific Data Format (ASDF) files, and JSON files. FITS format files contain the science data pixels, ASDF files contain the world coordinate system information, and the JSON files contain information regarding the way the science data is associated.

JWST data are FITS format files generated by the Data Management System (DMS) via the Science Data Processing (SDP) and the Calibration Pipeline. The telemetry data from the Recorded Science Data files that are received by DMS, is in the same binary format as stored on the JWST Solid State Recorder (SSR). These files will come in compressed packets that will be read by SDP to extract science data and relevant detector and exposure information.

The initial FITS header will contain keywords required by the FITS standard and keywords required for identification, naming of the files, data structure definition, and for the correct calibration of the science data by the calibration pipeline. These keywords are populated with telemetry packet headers and Science Image headers; proposal, planning, and scheduling information; spacecraft position; time conversions; pointing information; and select engineering parameters (JWST-STScI-002111).

The corresponding transformation from the detector positions to a world coordinate frame (ICRS and wavelength) for the science data are provided via distortion and spectral models stored in ASDF format extensions to the FITS file.

The JSON files provide with the information about how the data is associated. JWST data products can be divided into two main types of data; data products from single exposures produced during stage 1 and stage 2 of the calibration pipeline and data products that result from the combination of these exposures into a single one produced with stage 3. The difference between these two type of products can be easily told apart by their File Naming Conventions or their Header Keywords. Within an association, the exposure level data are combined into what we call stage 3 and level 4 data. Within these two categories, there are different ways in which a set or subset of exposures is combined, each of these corresponding to a unique association. The way in which these are combined is determined by the information coming from the Astronome Proposal Tool; matching the proposer organization of the observations via dithers, mosaics, and any other special requirement associating the data into a single set.

Not only the science data exposures will be part of an association. When considered necessary, an association will also include data that needs to be kept together for calibration purposes or post-calibration analysis by experienced observers.

File Naming Conventions and Data Products

JWST Data has unique names according to their relationship, within a proposal, with the observation mode and the members of an exposure and target. For a given association or exposure, the data also includes extension names according to the type of product.

Introduction

The naming convention for calibration products provides information about the program, exposure, instrument mode, as well as stages of calibration. At each stage of the calibration pipeline, the input data for a step is the calibration product produced by previous stages. In some cases, a given step of the calibration pipeline can produce other products that will not be used in subsequent steps and produced only when explicitly requested by the user. These are referred to as intermediate or optional products and are not stored in MAST.

Exposure file names

JWST science data files are FITS format files that in their first stages of processing contain only the values from a single exposure for a single detector. In the file name, a detector is defined to be a single Sensor Chip Assembly (SCA). These files are processed and combined with other exposures in stage 3 of the calibration pipeline.

The names of the exposure level data is constructed with information from the science header of the exposure, allowing users to map it to the observation in ther corresponding APT files:
jw<PPPPP><OOO><VVV>_<GGSAA>_<EEEEE>_<detector >_<suffix> .fits

where

- PPPPP = Program ID
- OOO = Observation number
- VVV = Visit number
- GG = Visit Group
- S = Parallel Sequence id (1 prime, 2-5 parallel)
- AA = Activity number (base 36)
- EEEEE = Exposure number
- <detector> = detector name; e.g. nrca1, nrclong, mirimage
- <sufix> = product type; e.g. uncal, rate

For example, the filename for a long wavelength uncalibrated NIRCam exposure, taken with module A, would take the form "jw80500012009_01101_00012_nrcalong_uncal.fits."

The observation number from the APT proposal entry is preserved as the observation number in the science telemetry, making it easier for the proposer to determine the observation that gave rise to an individual exposure. Since the exposure number resets with each new activity, the visit group, parallel sequence ID, and activity are included to make it possible to trace the exposure to the commanding request. All the products of the calibration pipeline that derive from this exposure have this naming convention, with the file name suffix distinguishing the different file products in the data set.

Stage 1 and stage 2 data products

The values for the suffix field within the file name, correspond to the different stages of the calibration pipeline and are as defined in the Algorithms Documentation. The input and output from each of stage 1 and stage 2 of the calibration pipeline steps are provided in Table 1.

Table 1. Standard and optional products for stage 1 and stage 2

Stage	Module	Input data suffix	Output product suffix	Units	Description
stage 0			_uncal	DN	Raw 4D (ncols x nrows x ngroups x nints) exposure data. Used as input to calwebb_detector1.
stage 1	calwebb_detector1	_uncal	_rateints	DN/s	3-D slope images for each integration. Each plane of the 3D cube is the result of an individual integration. Typically only of interest and used for time-series observations. Result of ramp fitting.
			_rate	DN/s	2-D slope images derived after averaging of all countrate integrations within the exposure.
			_ramp	DN	Optional-partial calibrated product. It is a corrected 4D exposure data. Processed up through jump detection step.
			_fitopt	(various)	Optional product with fitting results from ramp_fit step

			_trapsfilled	DN	Optional product. Charge trap state product produced by the persistence correction step.
			_dark	DN	Optional product. This is a 4D corrected dark exposure data. Processed up to dark subtraction step.
stage 2 (imaging)	calwebb_image2	_rate	_cal	DN/s	Calibrated slope images per exposure
			_i2d	DN/s	Rectified calibrated slope images per exposure.
			_bsub	DN/s	Background-subtracted image
		_rateints	_calints	DN/sec	Calibrated slope images for each integration
			_bsubints	DN/s	Per integration background-subtracted image
stage 2 (spectroscopy)	calwebb_spec2	_rate	_cal	DN/s (non-IFU) mJy/arcsec2 (IFU)	2D fully calibrated data. Data are averaged over all integrations.
			_s3d	mJy/arcsec2	IFU: 3-D spectral cubes per exposure
			_s2d	DN/s	non-IFU: 2-D rectified (drizzled) slope image per exposure
			_x1d	(various)	Extracted 1-D spectra over all integrations
			_bsub	DN/s	2D background-subtracted spectral data. Data are averaged over all integrations.
		_cal	_annn _crf		stage 3 product. Copy of _cal with DQ array updated by outlier_detection step
		_rateints	_x1dints	(various)	Extracted 1-D spectra per integration

			_calints	DN/s	Calibrated slope images for each integration
			_flat		Optional output for NIRSpec data. Provided the extracted and interpolated flat fields.
			_bsubints	DN/s	3D background-subtracted spectral data. Each plane of 3D cube is result for individual integrations.
		_calints	_annn _crfints		stage 3 product. Copy of _cal with DQ array updated by outlier_detection step

Association file names

Associations capture the relationship between exposures that are to be combined by design to form a single product. Association data products result from the combination of multiple exposures like dithers or mosaics and are generated by the stage 3 of the calibration pipeline. These products will have a file name that includes the astronomical source and observation parameters that would help users to map these with the proposed observation.

The format for the file name of an association has all alphabetic characters in lower case, underscores are only used to delineate between major fields, and dashes can be used as separators for optional fields. Just as for Stage 2, the suffix distinguishes the different file products of Stage 3 of the calibration pipeline:

jw<PPPPP>-<AC_ID>_x[<“t”TargetID | “s”SourceID>
](-<"epoch”X>)_<science_instrument>_<optical_elements>(-<subarray>)_ <suffix>(-<ACT_ID>).fits

where

- PPPPP = is the program ID
- AC_ID = association candidate (AC) ID which can take the following values
  - "o"oo0
  - "c"cccc
  - "a"aaaa
  - "r"rrrr
- “t”TargetID | “s”SourceID, either TargetID or SourceID must be present
  - TargetID = The standard target name is a 3 digit Target ID (usually imaging target)
  - SourceID = The standard source name is a 5 digit Source ID (usually spectral MSA target)
- "epoch”X = The text "epoch" followed by a single digit indicating the epoch number
- science_instrument = The designator for the science instrument, i.e. nircam, miri, etc.
- optical_elements> = A single or "-" separated list of optical elements alphabetically (filters).
- subarray = Optional indicator of the subarray.

- suffix = The product type suffix. See Link to the JCCWG confluence page listing product types.

In this case, the EPOCH is an optional parameter that is used for observations of the same target that are taken at different epochs.

Stage 3 data products

Table 2 is the Input exposure level data coming from stage 2 and their stage 3 products separated by the modules as described in the Stages of Processing article.

Table 2. Standard and optional products for stage 3 of the calibration pipeline

Data stage	Module	Input suffix	Description	Output suffix		Description
stage 3 imaging	calwebb_image3	_cal	All _cal files in an association	_i2d		2D rectified and combined image.
				_cat		Source catalog, in .ecsv format
				_crf	DN/s	Copy of _cal with DQ array updated by outlier_detection step
				_segm		Segmentation map
stage 3 spectroscopy	calwebb_spec3	_cal	All _cal files in an association	_crf	DN/s	Copy of _cal with DQ array updated by outlier_detection step
				_x1d	flux	Extracted 1-D spectra from combined exposures.
				_s3d		3D rectified and combined spectral cube. For IFU modes and spectral mapping observations.
				_s2d		non-IFU: 2D rectified and combined spectral image.
		_calint	All _calint files in an association	_crfints		Same as _calints but with updated CR flags

				_x1dints	net, flux	Extracted 1-D spectra per integration
stage 3 coronagraphic	calwebb_coron3	_calints	All _cal files in an association	_crf		Copy of _cal with DQ array updated by outlier_detection step
				_psfstack		3D stack of PSF images. Output of stack_refs step. One _psfstack product per associated target.
				_i2d		2D combined and resampled science target image. Output of resample step. One _i2d product per associated target
				_crfints		Same as _calints but with updated CR flags
				_annn_psfsub		psf subtracted
				_annn_psfalign		4D array of PSF images aligned to individual integrations of science target images. Output of align_refs step. One _psfalign product per target input.
stage 3 Time Series Observations	calwebb_tso3	_cal	All_cal files in an association	_s2d		2D resampled+combined spectroscopic image 1 product per source
				_crf	DN/s	Copy of _cal with DQ array updated by outlier_detection step
				_phot	mag	Imaging photometry, in ecsv format.
				_whtlt	.	Spectroscopic white-light curve, in ecsv format.

		_calints	All _calint files in an association	_crfints	DN/s	3D calibrated data with DQ array updated by outlier_detection step. Copy of _calints input.
stage 3 Aperture Masking Interferometry	calwebb_ami3	_cal	All_cal files in an association	_annn_ami		Output of ami_analize step. One per exposure
				_amiavg		Table of averaged fringe parameters for PSF reference or science target exposures. Output of ami_average step.
				_aminorm		Final single product for the target corrected by the calibrator

Header Keywords and Relationships

FITS Header Keywords provide information about the observation, type of data and post-observation processing. Full information for these can be found in the JWST Keyword Dictionary.

Introduction

All JWST data products are FITS files. The FITS header keywords of all the data files store important information characterizing the observations, telemetry received during the observations, and information that relates to the post-observation processing of the data. Following FITS conventions, each keyword is no longer than eight characters and these have values that can be integer, real (floating-point), or a character string. There are several keywords that are common to all JWST data while others are instrument specific. Knowledge of the keywords and where to find them is an important first step in understanding your data. By examining your file header, using fv, ds9, python's astropy.fits.getvalue function, you will find detailed information about your data including:

- Coordinates of the target, program number, and other observation identifiers.
- Date and time of observation including start, end and mid exposure times
- Exposure parameters information such as the instrument configuration information (detector, filter, subarray names and subarray positions)
- Readout definition parameters such as readpatt, nints, ngroups, and groupgap.
- Exposure-specific information such as more detailed timing and world coordinate system information.
- Calibration information such as the calibration switches and reference files used by the pipeline

Some of the keywords are only relevant to spectroscopic or imaging data, target acquisitions, and for specific instruments. Accordingly, users will find that the header on a particular file contains a unique combination of keywords appropriate for the type of observation. Long definitions for the keywords can also be accessed from the JWST Keyword Dictionary, which provides detailed explanations of the units, source, and many other attributes.

Keywords that deal with a particular topic, such as the program information or target information are grouped together logically throughout the headers. The sample of a data header below shows the keywords and groupings. These keywords can originate from the Proposal and Planning System (PPS), Observatory Status File (OSF), Science and Operations Center Project Reference Database (PRD), Science Data Processing (SDP), or Calibration Software, the source along with the FITS header extension where these can be found and valid values are provided in the source attribute of the JWST Keyword Dictionary.

Sample Header

```
SIMPLE = T / Data conform to FITS standard
BITPIX = 8 / bits per data value
NAXIS = 0 / number of data array dimensions
```

```
EXTEND = T / file may contain standard extensions
DATE = '2016-11-09T11:20:19' / [yyyy-mm-ddThh:mm:ss.ssssss] UTC date file creation
ORIGIN = 'STSCI ' / institution responsible for creating FITS file
TIMESYS = 'UTC ' / principal time system for time-related keywords
FILENAME= 'jw98012001001_0210E_00001_mirimage_uncal.fits' / Name of the file
FILETYPE= 'uncalibrated' / Type of data in the file
SDP_VER = 'UNKNOWN ' / data processing software version number
PRD_VER = 'PRDDEVSOC-D-012' / S&OC PRD version number used in data processing
TELESCOP= 'JWST ' / telescope used to acquire the data
HGA_MOVE= F / High Gain Antenna moved during data collection
RADESYS = 'ICRS ' / Name of the coordinate reference frame

COMMENT / Program information
TITLE = 'MIRI IMG-RAD-11B' / Proposal title
PI_NAME = 'N/A ' / Principal investigator name
CATEGORY= 'ENG ' / Program category
SUBCAT = '' / Program sub-category
SCICAT = '' / Science category assigned during TAC process
CONT_ID = 0 / Continuation of previous program

COMMENT / Observation identifiers
DATE-OBS= '2016-01-24' / [yyyy-mm-dd] UTC date at start of exposure
TIME-OBS= '07:47:22.005' / UTC time at the start of the exposure
OBS_ID = 'V98012001001P000000000210E' / Programmatic observation identifier
VISIT_ID= '98012001001' / Visit identifier
PROGRAM = '98012 ' / Program number
OBSERVTN= '001 ' / Observation number
VISIT = '001 ' / Visit number
VISITGRP= '02 ' / Visit group identifier
SEQ_ID = '1 ' / Parallel sequence identifier
ACT_ID = '0E ' / Activity identifier
EXPOSURE= '1 ' / Exposure request number

COMMENT / Visit information
TEMPLATE= 'MIRI External Flat' / Observation template used
OBSLABEL= 'IMAGER FLATS LAMP-ON ONLY' / Proposer label for the observation
VISITYPE= 'PRIME ' / Visit type
VSTSTART= '2016-01-24 06:59:44.0860000' / UTC visit start time
WFSVISIT= '' / Wavefront sensing and control visit indicator
VISITSTA= 'SUCCESSFUL' / Status of a visit
NEXPOSUR= 16 / Total number of planned exposures in visit
INTARGET= T / At least one exposure in visit is internal T
ARGOOPP= F / Visit scheduled as target of opportunity
```

```
COMMENT / Target information
TARGPROP= '' / Proposer's name for the target
TARGNAME= '' / Standard astronomical catalog name for target
TARGTYPE= 'FIXED ' / Type of target (fixed, moving, generic)
TARG_RA = 0.0 / Target RA at mid time of exposure
TARG_DEC= 0.0 / Target Dec at mid time of exposure
TARGURA = 0.0 / Target RA uncertainty
TARGUDEC= 0.0 / Target Dec uncertainty
PROP_RA = 0.0 / Proposer specified RA J2000 for the target
PROP_DEC= 0.0 / Proposer specified Dec J2000 for the target
PROPEPOC= 'Jan 1 2000 12:00AM' / Proposer specified epoch for proper motion

COMMENT /Exposure parameters
INSTRUME= 'MIRI ' / Instrument used to acquire the data
DETECTOR= 'MIRIMAGE' / Name of detector used to acquire the data
LAMP = '' / Internal lamp state
FILTER = 'F1140C ' / Name of the filter element used
CORONMSK= 'NULL ' / coronagraph mask used
CCCSTATE= 'OPEN ' / Contamination control cover state

COMMENT /Exposure information
PNTG_SEQ= 1 / Pointing sequence number
EXPCOUNT= 13 / Running count of exposures in visit
EXP_TYPE= 'MIR_FLATIMAGE' / Type of data in the exposure

COMMENT / Exposure times
EXPSTART= 57411.3245602429 / UTC exposure start time
EXPMID = 57411.32536320594 / UTC exposure mid time
EXPEND = 57411.32616616898 / UTC exposure end time
READPATT= 'FAST ' / Readout pattern
NINTS = 5 / Number of integrations in exposure
NGROUPS = 10 / Number of groups in integration
NFRAMES = 1 / Number of frames per group
GROUPGAP= 0 / Number of frames dropped between groups
NSAMPLES= 1 / Number of A/D samples per pixel
TSAMPLE = 10.0 / Time between samples (microsec)
TFRAME = 2.77504 / Time between frames (sec)
TGROUP = 2.77504 / Time between groups (sec)
EFFINTTM= 27.7504 / Effective integration time (sec)
EFFEXPTM= 139.25 / Effective exposure time (sec)
CHRGTIME= 0.0 / Charge accumulation time per integration (sec)
DURATION= 138.752 / Total duration of exposure (sec)
```

```
NRSTSTRT= 0 / Number of resets at start of exposure
NRESETS = 0 / Number of resets between integrations
ZEROFRAM= F / Zero frame was downlinked separately
DATAPROB= F / Science telemetry indicated a problem
SCA_NUM = 0 / Sensor Chip Assembly number
DATAMODE= 1 / post-processing method used in FPAP
COMPRSSD= F / data compressed on-board (T/F)
FRMDIVSR= 1 / Divisor applied to frame-averaged groups
SUBARRAY= 'FULL ' / Subarray used
SUBSTRT1= 1 / Starting pixel in axis 1 direction
SUBSTRT2= 1 / Starting pixel in axis 2 direction
SUBSIZE1= 1032 / Number of pixels in axis 1 direction
SUBSIZE2= 1024 / Number of pixels in axis 2 direction
FASTAXIS= 1 / Fast readout axis direction
SLOWAXIS= 2 / Slow readout axis direction
PATTTYPE= 'NONE ' / Primary dither pattern type
PATT_NUM= 0 / Position number in primary pattern
PATTSTRT= 0 / Starting point in pattern
NUMDTHPT= 0 / Total number of point in pattern
SUBPXNUM= 0 / Subpixel sampling pattern number
PATTSIZE= '' / Primary dither pattern size
SUBPIXEL= T / Subpixel sampling performed
XOFFSET = 0.0 / x offset from pattern starting position
YOFFSET = 0.0 / y offset from pattern starting position
COORDSYS= '' / Ephemeris coordinate system
EPH_TIME= 57143 / UTC time from ephemeris start time (sec)
JWST_X = -1436696.08564455 / X spatial coordinate of JWST (km)
JWST_Y = -713108.747788469 / Y spatial coordinate of JWST (km)
JWST_Z = -606536.187076095 / Z spatial coordinate of JWST (km)
JWST_DX = 0.0288294864316793 / X component of JWST velocity (km/sec)
JWST_DY = -0.0465945808751849 / Y component of JWST velocity (km/sec)
JWST_DZ = 0.0174011518182956 / Z component of JWST velocity (km/sec)
APERNAME= '' / PRD science aperture used
PA_APER = 0.0 / Position angle of aperture used (deg)
DVA_RA = -1.7189778168302E-07 / Velocity aberration correction RA offset (rad)
DVA_DEC = 1.23020859538226E-07 / Velocity aberration correction Dec offset (rad)
VA_SCALE= 1.000000060550096 / Velocity aberration scale factor
BARTDELT= 0.0 / Barycentric time correction
BSTRTIME= 0.0 / Barycentric exposure start time
BENDTIME= 0.0 / Barycentric exposure end time
BMIDTIME= 0.0 / Barycentric exposure mid time
HELIDELT= 0.0 / Heliocentric time correction
HSTRTIME= 0.0 / Heliocentric exposure start time
```

```
HENDTIME= 0.0 / Heliocentric exposure end time
HMIDTIME= 0.0 / Heliocentric exposure mid time
V2_REF = 0.0 / V2 coordinate of the reference pixel (SIAF:V2Re
V3_REF = 0.0 / V3 coordinate of the reference pixel (SIAF:V3Re
DEC_V1 = -38.74923877352641 / [deg] Dec of telescope V1 axis
NEXTEND = 3 / Number of file extensions
WCSAXES = 2 / number of World Coordinate System axes
CRPIX1 = 1024.0 / Axis 1 coordinate of the reference pixel in the
CRPIX2 = 1024.0 / Axis 2 coordinate of the reference pixel in the
CRVAL1 = 348.9278676961044 / First axis value at the reference pixel (RA in
CRVAL2 = -38.74923877352641 / Second axis value at the reference pixel (RA in
CTYPE1 = 'RA---TAN' / First axis coordinate type
CTYPE2 = 'DEC--TAN' / Second axis coordinate type
CUNIT1 = 'deg ' / units for first axis
CUNIT2 = 'deg ' / units for second axis
CDELT1 = 0.11 / first axis increment per pixel, increasing east
CDELT2 = 0.11 / Second axis increment per pixel, increasing nor
PC1_1 = 0.6404219811118805 / linear transformation matrix element cos(theta)
PC1_2 = -0.7680232327922992 / linear transformation matrix element -sin(theta
PC2_1 = 0.7680232327922992 / linear transformation matrix element sin(theta)
PC2_2 = 0.6404219811118805 / linear transformation matrix element cos(theta)
S_REGION= '' / spatial extent of the observation, footprint
RA_REF = 348.9278676961044 / Right Ascension of the reference point (deg)
DEC_REF = -38.74923877352641 / Declination of the reference point (deg)
ROLL_REF= 354.9135542146994 / Telescope roll angle of V3 North over East at t
VPARITY = 1 / Parity (sense) of aperture settings (1, -1)
V3I_YANG= 0.0 / Direction angle in V3 (Y)
GS_ORDER= 0 / index of guide star within list of selected gui
GSSTRTTM= '1999-01-01 00:00:00' / UTC time when guide star activity started
GSENDTIM= '1999-01-01 00:00:00' / UTC time when guide star activity completed
GDSTARID= '' / guide star identifier
GS_RA = 0.0 / guide star right ascension
GS_DEC = 0.0 / guide star declination
GS_URA = 0.0 / guide star right ascension uncertainty
GS_UDEC = 0.0 / guide star declination uncertainty
GS_MAG = 0.0 / guide star magnitude
GS_UMAG = 0.0 / guide star magnitude uncertainty
PCS_MODE= 'NONE ' / Pointing Control System mode
GSCENTX = 0.0 / guide star centroid x postion in the FGS ideal
GSCENTY = 0.0 / guide star centroid y postion in the FGS ideal
JITTERMS= 0.0 / RMS jitter over the exposure (arcsec).
VISITEND= '2016-01-24 07:58:05.1420000' / Observatory UTC time when the visit en
```

```
RA_V1 = 348.9278676961044
PA_V3 = 50.17670723125901
END
```

Understanding Associations

JWST associations are produced by combining science exposures using a set of predefined rules that depend on the type of data and observation.

Introduction

Associations provide the relationships between multiple exposures and provide the user with the means to identify a set of exposures that belong together and may be dependent upon one another. Associations allow for the data to be calibrated, archived, retrieved, and reprocessed as a group rather than as individual objects and allow the user to combine single, multiple observations or even different programs. Finally, associations capture the relationship between exposures and higher level data products.

In order to capture a list of exposures that could potentially form an association and provide relevant information about those exposures, DMS first generates an association pool. These are simple ASCII files that contain the metadata for all the data for a given proposal. These pools are then used by the association generator that runs within the calibration software to generate association definitions in a JSON format. Based on the association table content, the calibration software creates the associated data products.

The basic association that the pipeline generates are combinations of mosaics/dithers for a single observation. Higher level products are built by associating data outside routine science data processing pipeline; these might include multiple observations from a single target in a program or large mosaics similar to HLA for HST data.

Components of an association

An association can contain data products, related files, and contemporaneous calibration files.

- Data products are multiple exposures from a science instrument that are taken as part of a dither, a mosaic, a coronagraphic or AMI image, a time series observation, or moving targets.
- Related data files that provide support to the science data are included in an association. Supporting data files include jitter data, target acquisition images, background images and confirmation images.
- Contemporaneous calibration files are exposures used to calibrate the science data, such as wavelength calibrations, lamp data, or flat fields, which are executed in the same time frame as the science exposures.

Association Pool

There is a separate association for each program, and each target within a program if it can be determined that multiple targets in a program are not related. Each association will follow the same format. The pool of data that is considered to construct an association within a specific program is called association pool and it will include:

- All observations from the same target in a given program
- Observations from multiple science instruments of a given target within the same program
- Different filters for the same target within the same program
- Exposures from linked observations within the same program
- Calibration exposures can be members of more than one association pool

Association generator

When all of the exposures for an observation, or set of observations, have been collected, an association generator will determine which exposures are needed for the stage 3 (and in some instances stage 2) data products and will output an association table that documents the content of the association. Multiple association tables can be generated from a single association pool. In DMS, associations are created by the association generator which, based on rules, classifies the data into one or more associations. Users should not need to run the generator; instead, it is expected they will edit the existing association that accompanies the user's JWST data.

An association table is a JSON formatted file that includes the list of all data related that might or might not be combined into a single image. An example of the format for this file is below.


```
{
"asn_rule": "candidate_Asn_Image",
"asn_id": "o001",
"target": "1",
"code_version": "0.7.1.beta4",
"version_id": "20170703T120130",
"asn_pool": "jw96090_20170703T120130_pool",
"degraded_status": "No known degraded exposures in association.",
"program": "96090",
"products": [
{
"name": "jw96090-o001_t001_nircam_f322w2-f323n",
"members": [
{
"exposerr": "null",
"exptype": "science",
"expname": "jw96090001004_03101_00001_nrcalong_cal.fits",
"asn_candidate": "('o001', 'observation')"
},
```

```
{
"exposerr": "null",
"exptype": "science",
"expname": "jw96090001003_03101_00001_nrcalong_cal.fits",
"asn_candidate": "('o001', 'observation')"
},
{
"exposerr": "null",
"exptype": "science",
"expname": "jw96090001002_03101_00001_nrcalong_cal.fits",
"asn_candidate": "('o001', 'observation')"
},
{
"exposerr": "null",
"exptype": "science",
"expname": "jw96090001001_03101_00001_nrcalong_cal.fits",
"asn_candidate": "('o001', 'observation')"
},
{
"exposerr": "null",
"exptype": "science",
"expname": "jw96090001003_03101_00001_nrcblong_cal.fits",
"asn_candidate": "('o001', 'observation')"
},
{
"exposerr": "null",
"exptype": "science",
"expname": "jw96090001004_03101_00001_nrcblong_cal.fits",
"asn_candidate": "('o001', 'observation')"
},
{
"exposerr": "null",
"exptype": "science",
"expname": "jw96090001001_03101_00001_nrcblong_cal.fits",
"asn_candidate": "('o001', 'observation')"
},
{
"exposerr": "null",
"exptype": "science",
"expname": "jw96090001002_03101_00001_nrcblong_cal.fits",
"asn_candidate": "('o001', 'observation')"
}
]
```

```
}
],
"asn_type": "image3",
"constraints": "Constraints:\n target: 1\n opt_elem2: f323n\n opt_elem: f322w2\n
```

```
exp_type: nrc\\_image\n instrument: nircam\n program: 96090\n wfsvisit: Is
Invalid\n asn_candidate: \\(\\'o001\\'\\,\\ \\'observation\\'\\)"
}
```

Association names

Association names have to be unique and allow for different possible types of data to be associated. Associations produced in stage 3 of the calibration pipeline have the following format:

```
jw<ProgramID>-<AC_ID>_<target|source
ID>[-<"epoch”X>]_<science_instrument>_<optical_elements>(-<subarray>)_
<product_type>(-<ACT_ID>).fits
```

where

- <ProgramID> or <ppppp> is the program identifier,
- <AC_ID> is the association candidate (AC) ID. This AC ID table is a byproduct of the APT. There are four types of association candidates:
    - <"o"ooo> can be o001 - o999 and are for associations constructed directly from all observations in an observation folder in APT.
    - <"c"cccc> can be c1000 - c2999 and are for candidate associations constructed from linked observations within a program via APT
    - <"a"aaaa> can be a3000 - a4999 and are for archive associations, constructed from linked observations within a program but not explicit via APT
    - <"r"rrrr> can be r5000 - r9999 and are for reserved associations for future use. These include associations that do not fall within the above types; i.e. high-level products.
- <target|source ID> one of these should be present
    - <tTTT> is a three-digit target ID. Usually for imaging targets.
    - <sSSSSS> is a five digit source ID. Usually for spectral targets.
- <"epoch”X> is the text "epoch" followed by a single digit indicating the epoch number. This is an optional parameter.
- <science_instrument> is the science instrument; e.g. nircam, miri, nirspec, etc.
- <optical_elements> is a list of optical elements separated by "-"; e.g grating-filter
- <subarray> this is an optional parameter for subarrays.
- <product_type> is the suffix for the product type. See File Naming Conventions and Data Products for a listing product types.
- <ACT_ID> is a two-digit number indicating the activity ID

Underscores are used to separate fields within the file name. Dashes are used to separate values within fields. An example for an imaging association would look like:

```
jw84140-o003_t001-epoch2_nircam_f090w_i2d.fits or
jw80500-o004_t002_nircalong_f356w_cal.fits
where we set observation ooo = 3 or 4, target number TTT = 001 or 002, and epoch =
2 on none
```

Type of associations

The following tables provide with a list of data types that might belong to an association; if these exist for that type of observation. Any background or PSF observation can belong to more than one association.

Table 1. How data is associated

Type of data		Association			
		Imaging	Spectroscopic	Coronographic	AMI
Target Acquisition		x	x	x	x
Astrometric confirmation images		x	x	x	
Singe target observation	All dither points	x	x	x	x
	All nodding points		x	x	
	All mosaic tiles within an observation folder	x	x		
	All mosaic tiles in different observation folders	x	x		

	In different orientation (target grouped via special requirements sequence observation non-interruptable)			x	x
Background observations		x	x		
Autowave calibration observations			x		
Autoflat calibration observations			x		
Confirmation Images			x		
Pre-image			x		
MSA Plan sources catalog			x		
Leak image			x		
PSF observation associated with the science target				x	

Outlier Detection Associations Products

The Outlier Detection step is part of stage 3 of the calibration pipeline; however, its products have the characteristics of those produced in stage 2. These are a copy of the _cal files produced in stage 2 but with the DQ array updated with flags for new outliers detected as part of the mosaicing or cube building step. The outliers are mainly due to cosmic rays that were not detected and handled as part of CALDETECTOR1.

Because these files are produced in stage 3, these have in their root names the information about the association ID from which these were derived. In this case, the name has the form

```
jw<PPPPP><OOO><VVV>_<GGSAA>_<EEEEE>_<detector >_<AC_ID>_<suffix> .fits
```

where all the fields, except for <AC-ID> and <suffix>, follow the same naming convention as those of the exposure file names produced in stage 2. The <AC_ID> parameter takes the value of the association ID that created them and can have the form of any of the association candidate types. These are referred as stage 2c products and these can be recreated as new data for a given association becomes available. In order to clearly distinguish these products from the original stage 2 counterparts, the <suffix> will be different.

Table 2. Stage 2c products suffix and associated stage 2 data

Stage 3 suffix	Type of data
_crf	2D calibrated data with DQ array updated by outlier_detection step. Copy of _cal input.
_crfints	3D calibrated data with DQ array updated by outlier_detection step. Copy of _calints input. One _annnn_crfints product per target _calints input (none for PSF inputs).

In this case, the _crf suffix stands for "cosmic-ray flags".

Table 3 shows some examples of stage 2c products. The first and second associations shown here are constructed from a two-point dither, 2 × 1 mosaic; which will generate four stage 2c products (eight if NINT >1), each with the DQ array updated after combining these observations. The third example is a mosaic-of-mosaics association that uses the two previously created associations as input. Note that in the latter we use association ID <"c"ccc> because we assume this is an association constructed from linked observations within a program via APT.; otherwise, we would have used <"a"aaa>.

Table 3. Example association and its stage 2c products

Association	Exposure	dither point	mosaic point
jw80500-o001_t001_nircam_f444w-clear_cal.fits	jw80500001001_01101_00001_nrcalong_cal.fits jw80500001001_01101_00001_nrcalong_o001_crf.fits (2c)	1	1
	jw80500001002_01101_00001_nrcalong_cal.fits jw80500001002_01101_00001_nrcalong_o001_crf.fits (2c)	2	1

	jw80500001003_01101_00001_nrcalong_cal.fits jw80500001003_01101_00001_nrcalong_o001_crf.fits (2c)	1	2
	jw80500001004_01101_00001_nrcalong_cal.fits jw80500001004_01101_00001_nrcalong_o0001_crf.fits (2c)	2	2
jw80500-o002_t001_nircam_ f444w-clear_cal.fits	jw80500003001_01101_00001_nrcalong_cal.fits jw80500003001_01101_00001_nrcalong_o002_crf.fits (2c)	1	1
	jw80500003002_01101_00001_nrcalong_cal.fits jw80500003002_01101_00001_nrcalong_o002_crf.fits (2c)	2	1
	jw80500003003_01101_00001_nrcalong_cal.fits jw80500003003_01101_00001_nrcalong_o002_crf.fits (2c)	1	2
	jw80500003004_01101_00001_nrcalong_cal.fits j w80500003004_01101_00001_nrcalong_o0002_crf.fits (2c)	2	2
jw80500-c001_t001_nircam_ f444w-clear_cal.fits (combines o001 and o002 above)	jw80500001001_01101_00001_nrcalong_cal.fits jw80500001001_01101_00001_nrcalong_c001_crf.fits (2c)	1	1
	jw80500001002_01101_00001_nrcalong_cal.fits jw80500001002_01101_00001_nrcalong_c001_crf.fits (2c)	2	1
	jw80500001003_01101_00001_nrcalong_cal.fits jw80500001003_01101_00001_nrcalong_c001_crf.fits (2c)	1	2

	jw80500001004_01101_00001_nrcalong_cal.fits jw80500001004_01101_00001_nrcalong_c0001_crf.fits (2c)	2	2
	jw80500003001_01101_00001_nrcalong_cal.fits jw80500003001_01101_00001_nrcalong_c001_crf.fits (2c)	1	1
	jw80500003002_01101_00001_nrcalong_cal.fits jw80500003002_01101_00001_nrcalong_c001_crf.fits (2c)	2	1
	jw80500003003_01101_00001_nrcalong_cal.fits jw80500003003_01101_00001_nrcalong_c001_crf.fits (2c)	1	2
	jw80500003004_01101_00001_nrcalong_cal.fits jw80500003043_01101_00001_nrcalong_c001_crf.fits (2c)	2	2

Note that in the bove example we are omitting the _calint and _crfint products that are created when NINTS > 1.

Time-series observations

These type of observations distinguish themselves by a large number of integrations that can result in very large stage 1 and stage 2 files; these can be 200 Gb in size or more. Transferring or trying to calibrate files of this size can be limiting to the user so DMS will break these files into more manageable files sizes documenting their relationship and order within the association, such that user can easily reconstruct their observations.

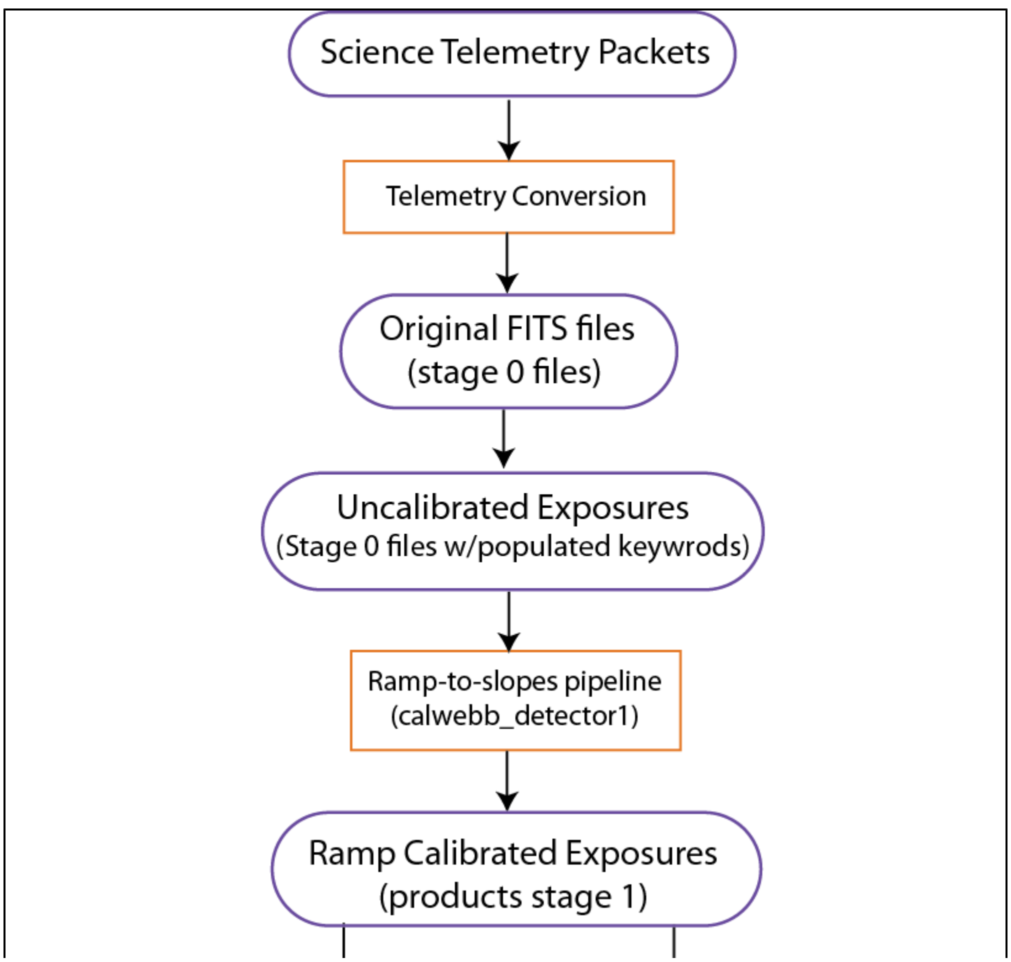
JWST Data Structure

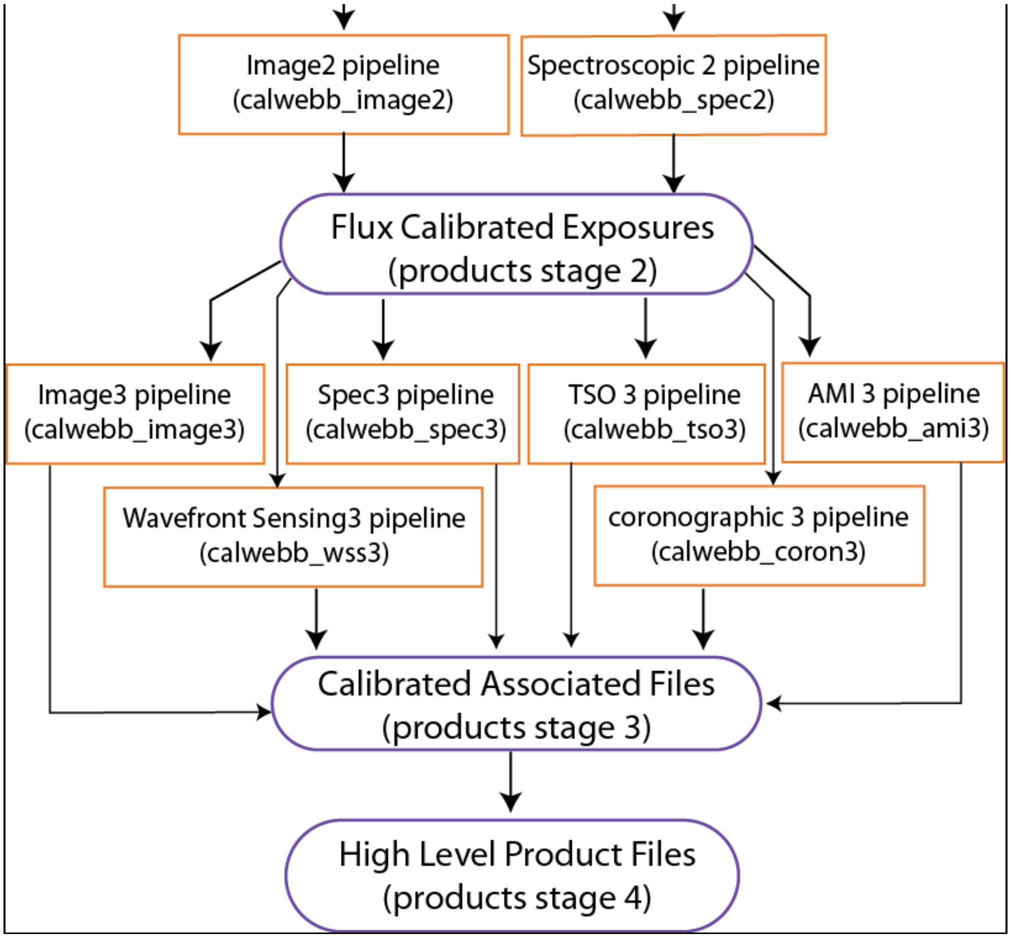
The JWST data files contain information about the science image or spectra; science proposal, planning, and scheduling information; spacecraft position; time conversions; pointing information; and select engineering parameters. Some of this information might change and the data format is updated as the data goes through the calibration steps of the pipeline. There are also optional files that contain extra information about the calibration that might be useful to users to further understand their observations.

Introduction

JWST data are FITS format files. Uncalibrated data (also known as stage 0 data files), as well as all data processed in the first two stages of the calibration pipeline, are science data FITS files which contain the pixel values for a single exposure from a single detector. The calibration pipeline in its stage 3, associates a set of these files into a single unit. Figure 1 shows the flow of the data from the telemetry packages to the most combined set of data.

Figure 1. Flow of data from the telemetry packages to the most combined set of data

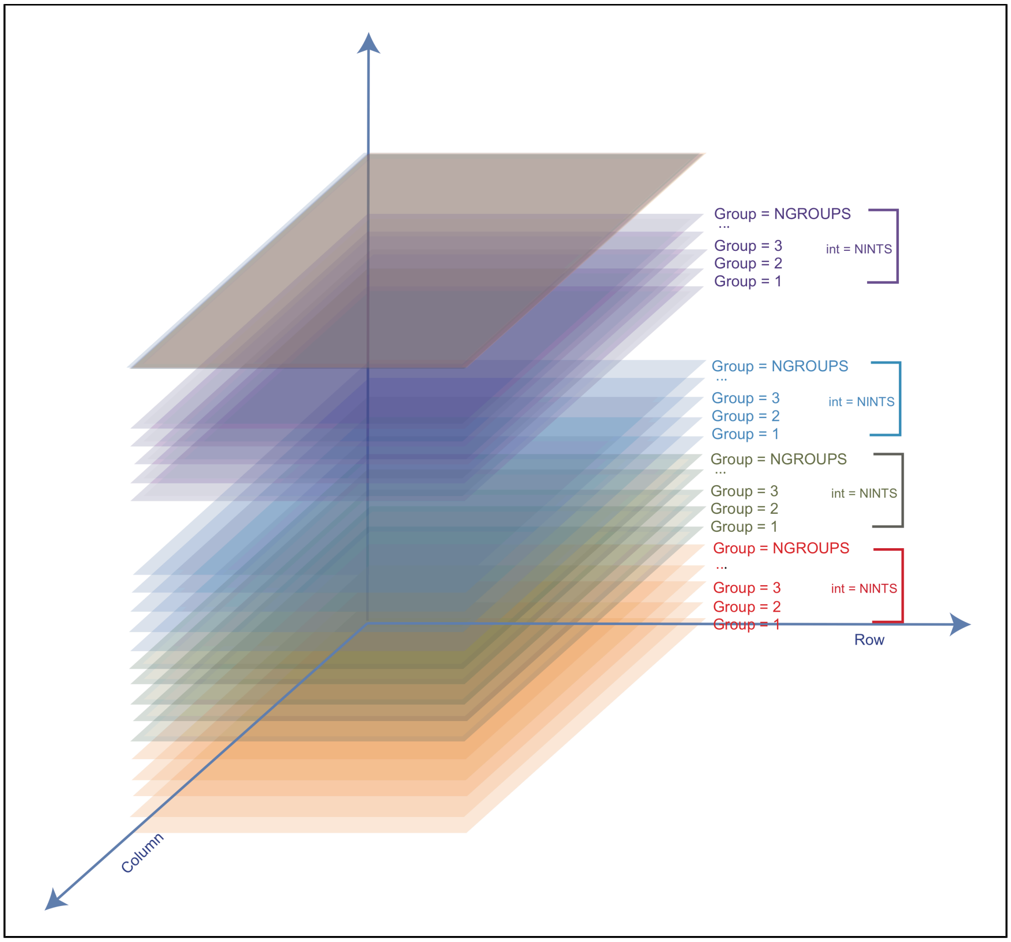




Stage 0 files (_uncal)

After telemetry conversion and coordinate transformation, Science Data Processing (SDP) will generate stage 0 FITS files with the pixel data by detector and exposure. SDP will format the science data into a 4D data array.

Figure 2. Science exposure data cube

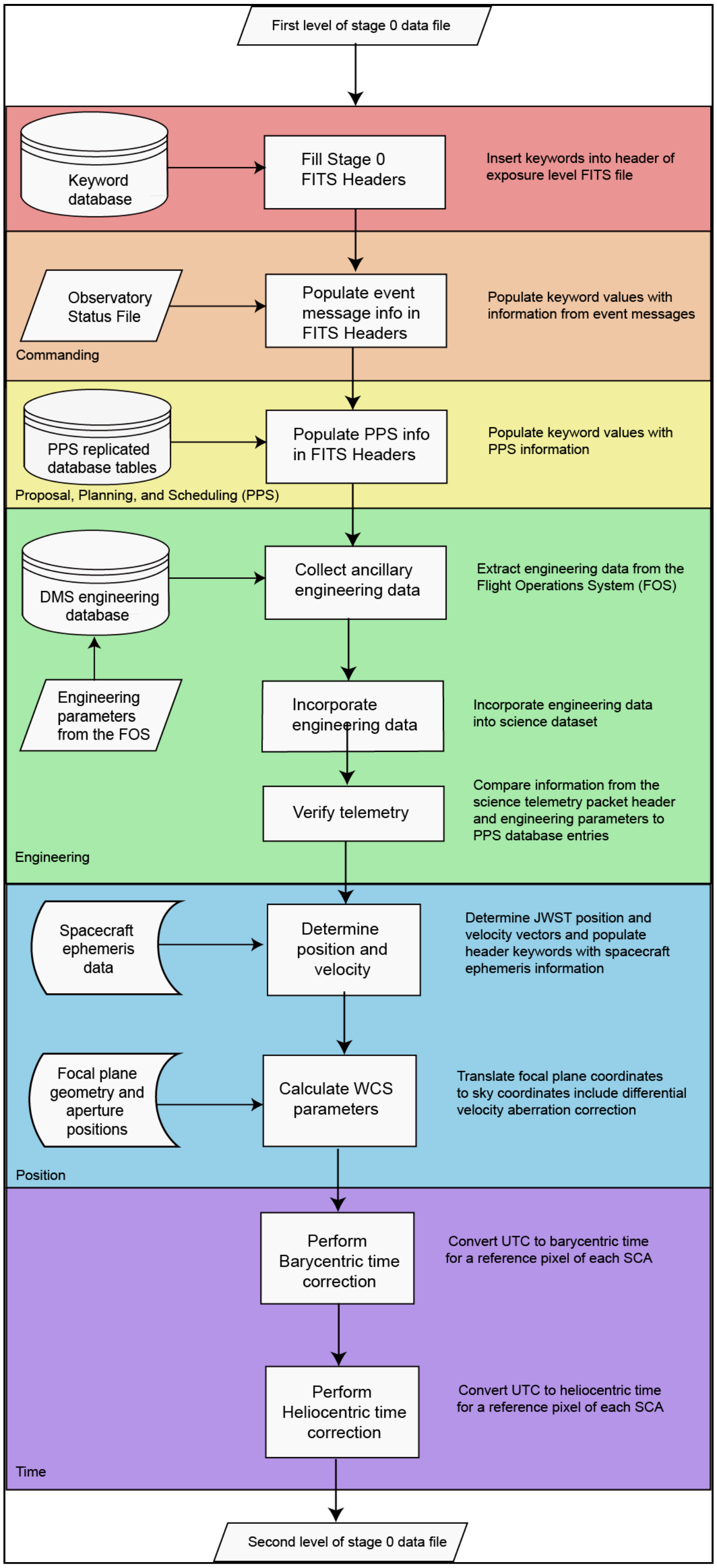


NAXIS1 = column, NAXIS2 = row, NAXIS3= Ngroups, NAXIS4 = Nints

The first two dimensions are the 2D science images, the dimension of the third axis is determined by the number of groups in the exposure, while the fourth axis corresponds to the number of integrations. All four dimensions will be used, even if NGROUPS = 1 or NINTS = 1. Each integration within an exposure is required to have the same number of groups. Due to up-the-ramp accumulation, pixel values increase between groups within an integration. The first frame of an integration may be read out separately, even if that frame is also averaged into the first group of the integration. A zero frame readout is identified with Group ID = 0. If there is a zero frame readout, there will be one for each integration. The zero frame will be stored in the data cube as the first image in each integration.

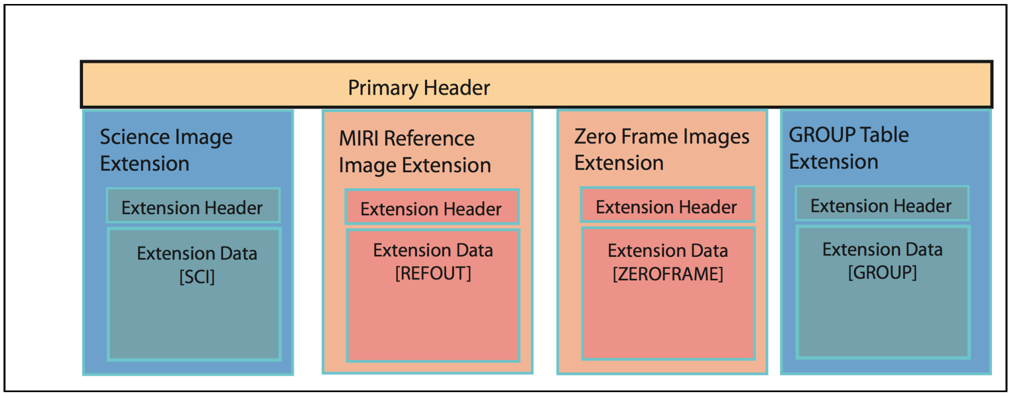
The FITS header of these files will have keywords required by the FITS standard and those extracted from the CCSDS telemetry packet headers and science image headers. The second phase of stage 0 of the processing of these files will add additional keywords/value pairs to the stage 0 files. These additional header data includes the proposal's planning and scheduling information; spacecraft position; time conversions; pointing information; and select engineering parameters. Figure 3 is a schematic representation of how the data from many sources (planing, scheduling, and observatory) is used by the Science Data Processing system to populate the FITS headers of the science data.

Figure 3. Source and usage of information by the Science Data Processing system to populate the headers of the science data.



At the completion of science data processing, the stage 0 (_uncal) data file will be ready for calibration. The overall file structure of the _uncal files is shown in the diagram below

Figure 4. Stage 0 (_uncal) file structure



The Primary Header Data Unit (HDU) contains keywords that apply to all subsequent extensions

- Science Data (SCI) extension – contains the uncalibrated pixel values in an image extension.
- Reference image (REFOUT) extension – contains reference pixels from the fifth output line read out simultaneously with the science data for MIRI.
- Zero frame images (ZEROFRAME) — contains the zero-frame data, if requested to be downlinked,
- Group Information (GROUP) extension – contains an ASCII table of group information extracted from the telemetry image header

From stage 0 to stage 3 data

When running the full calibration pipeline modules the default and optional output files are listed in the table of data products. All data products will have:

- a primary header unit
- one or several extensions, each with their own header unit

just as shown in the stage 0 product diagram above. The type number and type of extensions depend on the type of data product, and in some cases depending on the instrument. The type of products from the calibration pipeline can be classified into 3 main categories:

1. Default products produced by the calibration pipeline module and archived (default)
2. Optional products produced only when the calibration pipeline is run from the command line and the parameter to save that particular file is set to True in the particular step that produces it (optional)
3. The output from each of the calibration steps when each step is run independently rather than as a complete module such as calwebb_detector1. (intermediate)

Table 1 provides information for each of these products.

Output products

In here we provide with the complete list of default and optional products of the calibration pipeline and their format.

Table 1. Default and optional output products

Step	Module	Type	Product extension
		default	_uncal
Group Scale Correction	calwebb_detector1		_groupscalestep
Data Quality Initialization		intermediate	_dqinitstep
Saturation Detection		intermediate	_saturationstep
Superbias Subtraction		intermediate	_superbiasstep
Reference Pixel Correction		intermediate	_refpixstep
Linearity Correction		intermediate	_linearity
Persistence Correction		intermediate	_persistencestep
		optional	_output_pers
		default	_trapsfilled
Dark Current Subtraction		intermediate	_darkcurrentstep
		optional	
Jump Detection		intermediate	_jumpstep
Ramp Fitting		default	_rate
		default	_rateints
		optional	_ramp
Last Frame correction		intermediate (MIRI only)	_lastframestep
RSCD Correction		intermediate (MIRI only)	rscd_step.fits

Assign WCS		intermediate	_assignwcsstep
Background Image Subtraction		intermediate	_backgroundstep
		optional	_bsub
		optional	_bsubints
Flat Field Correction		intermediate	_flatstep
		optional	[_flat]
Photometric Correction		intermediate	_photomstep
Resample		default	_cal
		default	_calints
		default	_i2d
		defailt	_s3d
		default	_s2d
		default	_x1d
		default	_x1dints

File format

The format for each of the calibration products is provided below. These appear in alphabetic order.

assignwcsstep.fits

The output fits file for this step will have the same format as the ramp step

Primary Header: inserts calibration status keyword S_WCS and reference file recorded in R_CAMERA, R_COLLIMATOR, R_DISPERSER, R_DISTORTION, R_FORE, R_FPA, R_MSA, R_OTE, R_REGIONS, R_WAVELENGTHRANGE keywords

backgroundstep.fits

The output fits file for this step will have the same format as the ramp step

Primary Header: inserts calibration status keyword S_BKDS keyword

darkcurrentstep.fits

The output fits file for this step will have the same format as the dqinitstep step

Primary Header: inserts calibration status keyword S_DARK and reference file recorded in R_DARK keywords.

dqinitstep.fits

The output FITS file for this step will have the extensions listed in Table 2.

Table 2. File format for the dqinit file

No.	Name	Type	Cards	Dimensions	Format	Units
0	PRIMARY	PrimaryHDU	218	()		
1	SCI	ImageHDU	11	(ncols, nrows, ngroups, nints)	float32	DN
2	PIXELDQ	ImageHDU	10	(ncols, nrows)	int32 (rescales to unit32)	n/a
3	GROUPDQ	ImageHDU	10	(ncols, nrows, ngroups, nints)	uint8	n/a
4	ERR	ImageHDU	10	(ncols, nrows, ngroups, nints)	float32	DN
5	REFOUT	ImageHDU (MIRI only)	10	(258, 1024, ngroups, nints)	float32	DN
6	GROUP	BinTableHDU	37	6R x 13C	[I, I, I, J, I, 26A, I, I, I, I, 36A, D, D]	n/a
7	ASDF	ImageHDU	7	(4760,)	uint8	n/a

where the GROUP extension is a table with columns

```
name:
    ['integration_number', 'group_number', 'end_day',
'end_milliseconds', 'end_submilliseconds', 'group_end_time',
'number_of_columns', 'number_of_rows', 'number_of_gaps',
'completion_code_number', 'completion_code_text', 'bary_end_time',
'helio_end_time']
unit:
    ['', '', '', '', '', '', '', '', '', '', '', 'MJD', 'MJD']
```

and the ASDF extension is contained in every data product file coming out of the cal pipeline. It stores meta data for the product. The main constituents are the values of header keywords and, most importantly, all of the detailed data used by the World Coordinate System transformations. The metadata contained in the ASDF extension is not terribly useful at the file level but is used extensively within the calibration pipeline processing code.

Header input: inserts calibration status keyword S_GRPSCL and reference file recorded in R_MASK keyword.

flatfieldstep.fits

The output fits file for this step will have the same format as the ramp step

Primary Header: inserts calibration status keyword S_FLAT keyword and reference files recorded in R_FLAT keyword

groupscale.fits

The output FITS file for this step will have the extensions listed in Table 3

Table 3. File format for the groupscale file

No.	Name	Type	Cards	Dimensions	Format
0	PRIMARY	PrimaryHDU	42	()	
1	SCI	ImageHDU	309	(ncols, nrows, ngroups, nints)	float32
2	PIXELDQ	ImageHDU	10	(ncols, nrows)	int32 (rescales to unit32)
3	GROUPDQ	ImageHDU	10	(ncols, nrows, ngroups, nints)	uint8
4	ERR	ImageHDU	10	(ncols, nrows, ngroups, nints)	float32
5	REFOUT	ImageHDU (MIRI only)	10	(258, 1024, ngroups, nints)	float32
6	ASDF	ImageHDU	7	(20445,)	uint8

jumpstep.fits

The output fits file for this step will have the same format as the dqinitstep step

Primary Header: inserts calibration status keyword S_JUMP and reference file recorded in R_READNS, R_GAIN keyword.

lastframestep.fits

The output fits file for this step will have the same format as the dqinitstep step

Primary Header: inserts calibration status keyword S_LASTFR.

linearitystep.fits

The output fits file for this step will have the same format as the dqinitstep step

Primary Header: inserts calibration status keyword S_LINEAR and reference file recorded in R_LINEAR keyword.

persistencestep.fits

The output fits file for this step will have the same format as the dqinitstep step

Optional output file _output_pers has format

Primary Header: inserts calibration status keyword S_PERSIS and reference file recorded in R_TRAPDS, R_PERSAT, and R_TRAPAR. keywords.

photomstep.fits

The output FITS file for this step will have the extensions listed in Table 4

Table 4. File format for the photomstep file

No.	Name	Type	Cards	Dimensions	Format	
0	PRIMARY	PrimaryHDU	247	()		
1	SCI	ImageHDU	9	(ncols, nrows)	float32	
2	DQ	ImageHDU	10	(ncols, nrows)	int32	int32 (rescales to unit32)
3	ERR	ImageHDU	8	(ncols, nrows)	float32	
4	AREA	ImageHDU 8 (1024, 1024) float32	8	(ncols-refpix, nrows)	float32	
5	ASDF	ImageHDU	7	(5037,)	uint8	

Primary Header: inserts calibration status keyword S_PHOTOM keyword and reference files recorded in R_PHOTOM keyword

rampfit.fits

The output FITS file for this step will have the extensions listed in Table 5

Table 5. File format for the rampfit file

No.	Name	Type	Cards	Dimensions	Format	
0	PRIMARY	PrimaryHDU	247	()		
1	SCI	ImageHDU	9	(ncols, nrows)	float32	
2	DQ	ImageHDU	10	(ncols, nrows)	int32	int32 (rescales to unit32)
3	ERR	ImageHDU	8	(ncols, nrows)	float32	
4	ASDF	ImageHDU	7	(5037,)	uint8	

Primary Header: inserts calibration status keyword S_RAMP and reference file recorded in R_READNS keyword.

This step produces an optional product called ramp.fits with the extensions listed in Table 6

Table 6. File format for the ramp file

No.	Name	Type	Cards	Dimensions	Format
0	PRIMARY	PrimaryHDU	7	()	
1	SLOPE	ImageHDU	10	(ncols, nrows,ngroups, nints)	float32
2	SIGSLOPE	ImageHDU	10	(ncols, nrows, ngroups, nints)	float32
3	YINT	ImageHDU	10	(ncols, nrows, ngroups, nints)	float32
4	SIGYINT	ImageHDU	10	(ncols, nrows, ngroups, nints)	float32
5	PEDESTAL	ImageHDU	9	(ncols, nrows, nints)	float32
6	WEIGHTS	ImageHDU	10	(ncols, nrowsn, ngroups, nints)	float32
7	CRMAG	ImageHDU	10	(ncols, nrows,11,1)	float32
8	ASDF	ImageHDU	7	(1187,)	uint8

refpixstep.fits

The output fits file for this step will have the same format as the dqinitstep step

Primary Header: inserts calibration status keyword S_REFPIX

rscd_step.fits

MIRI only step.

The output fits file for this step will have the same format as the dqinitstep step

Primary Header: inserts calibration status keyword S_RSCD and reference file recorded in R_RSCD keyword

saturationstep.fits

The output fits file for this step will have the same format as the dqinitstep step

Primary Header: inserts calibration status keyword S_SATURA and reference file recorded in R_SATURA keyword.

superbiasstep.fits

The output fits file for this step will have the same format as the dqinitstep step

Primary Header: inserts calibration status keyword S_SUPERB and reference file recorded in R_SUPERB keyword.

trapsfilled.fits

The output FITS file for this step will have the extensions listed in Table 6

Table 6. File format for the trapsfilled file

No.	Name	Type	Cards	Dimensions	Format
0	PRIMARY	PrimaryHDU	141	()	
1	SCI	ImageHDU	9	(100,100,3)	float32
2	ASDF	ImageHDU	7	(2689,)	uint8

This is a default output of the Persistence Correction step.

References

JWST-STScI-00211, "DMS Level 1 and 2 Data Product Design"
JWST-STScI- 004078, " Design of Imaging Associations"
JWST technical documents

Working with FITS Files

JWST data is assembled into FITS files. Several software packages are available to manipulate and visualize JWST Data

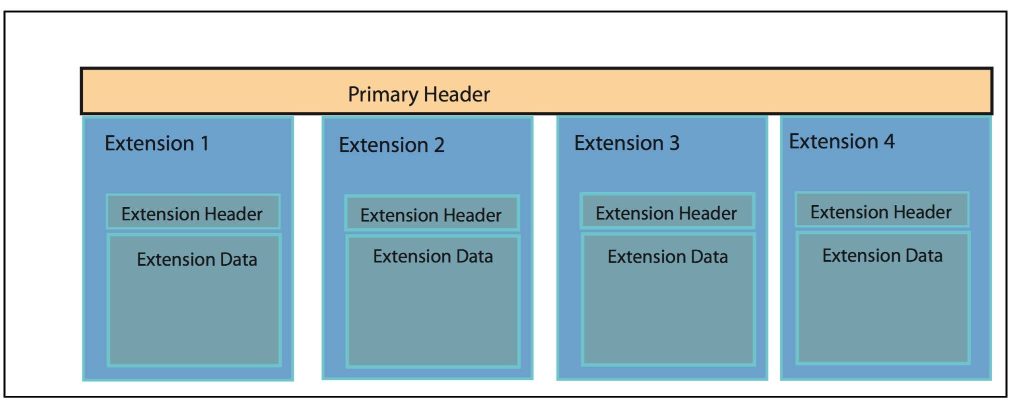
Introduction

The DMS pipeline automatically will process and calibrate all the data received from JWST and assembles them into a form suitable for most scientific analyses. JWST data will be made available to observers as files in multi-extension FITS format. JWST data can be manipulated with several different software packages. In this section, we introduce a few of the software options.

Multi-extension FITS format

Flexible Image Transport System (FITS) is a standard format for exchanging astronomical data, independent of the hardware platform and software environment.

FITS format files consist of a series of Header Data Units (HDUs), each containing two components: an ASCII text header and binary data. The header contains a series of keywords that describe the data in a particular HDU; the data component may immediately follow the header.
For JWST FITS data, the first HDU, or primary header, contains no data. The primary header may be followed by one or more HDUs called extensions. Extensions may take the form of images, binary tables, ASDF files, or ASCII text tables. The data type for each extension is recorded in the XTENSION header keyword. The figure below shows a schematic representation of such FITS file structure.



Each FITS extension header contains the required keyword XTENSION, which specifies the extension type and has one of the following values: IMAGE, BINTABLE, and TABLE, corresponding to an image, binary table, and ASCII or ASDF table, respectively.
Table 1. Extension information

HDU keyword	Description	Values	
XTENSION	Data type for extension	• IMAGE • BINTABLE (binary table) • TABLE (ASCII table)	
EXTNAME	Extension names that describe the type of data component for default outputs of the calibration pipeline	• SCI (science image) • ERR (error image) • DQ (data quality array) • PIXELDQ ("2-D pixel-dependent data quality flags) • GROUPDQ (4-D group-dependent data quality flags) • REFOUT (MIRI reference output) • ASDF (product metadata, includingWorkdCoordinateSystemtransforms • GROUP (Group metadata from telescope telemetry) • DQ_DEF (table with list of bit assignments for flag conditions used in the DQ array) • AREA (pixel area map) • EXTRACT1D (binary table with 1-D extracted spectra; part of the _x1d and _x1d products	
	Additional extension names for the optional output product from the ramp_fit step	• SLOPE (4-D arrays with slopes) • SIGSLOPE (4-D arrays with uncertainty in the y-intercept) • YINT (4-D arrays with y-intercept) • SIGYINT(4-D arrays with uncertainty in the y-intercept) • WEIGHTS (4-D arrays of fitting weights for each ramp interval of each pixel) • EXTRACT1D (binary table with 1-D extracted spectra; part of the _x1d and _x1d products	

Working with multi-extension FITS images and tables

Python

Python is used for astronomical data reduction applications. It is a freely available, general-purpose, dynamically-typed interactive language that provides modules for scientific programming and is used for astronomical data reduction application. These modules include:

- astropy package provides access to FITS files.

- numpy an IDL-style array manipulation facilities
- matplotlib plotting and image display package

Python is a very powerful language that is well suited to writing programs to solve many needs besides scientific analysis. Tutorials are available which illustrate the use of Python for interactive data analysis in astronomy (in much the same style as is now popular with IDL). The initial focus of these tutorials is the use of interactive tasks for the novice user. The more advanced tutorials focus on teaching the details of Python programming. The tutorials can be downloaded from:

http://www.scipy.org/Topical_Software. All the JWST Calibration pipelines are written in Python. More information on the use of Python to analyze JWST data can be obtained from JWST Post-Pipeline Data Analysis

IDL

IDL is an array-based, interactive programming language that provides many numerical analysis and visualization tools and is very popular in the astronomical community with many astronomers using it for their analysis of data. It can be obtained from ITT Visual Information Solutions for a fee. Libraries for reading astronomical FITS data are part of the freely available on the ASTRON library which has links to other IDL astronomy libraries.

Fortran and C

For those who wish to write their own Fortran or C applications use the FITSIO library for reading FITS files with Fortran and the CFITSIO library for C.

PyRAF/STADAS/TABLES

PyRAF is the python-based "Image Reduction and Analysis Facility" (IRAF) system that includes a selection of programs for general image processing and graphics, plus a number of programs for the reduction and analysis of optical and IR astronomy data.

The Space Telescope Science Data Analysis System (STSDAS), which is also layered on top of PyRAF, is part of the software tools offered at STScI and used mainly for HST data analysis.

The TABLES package sits alongside STSDAS and provides tools and libraries for working with tabular data. STSDAS requires TABLES, but one may use TABLES without STSDAS. Together, these two packages comprise STSDAS/TABLES.

For more information and support for IRAF, please refer to the NOAO pages at http://iraf.noao.edu/

> Other packages to work with FITS files are discussed in the JWST Post-Pipeline Data Analysis tools page.

ASDF Data

The Advanced Scientific Data Format (ASDF) is used to store models for axes manipulation and rules that serve to exchange information with scientist of between stages of the calibration pipeline. This format is used by some reference files and also in some of the science calibrated products.
ASDF is a human-readable, hierarchical metadata structure, made up of basic dynamic data types such as strings, numbers, lists, and mappings intended as an interchange format for delivering products from instruments to scientists or between scientists.

For JWST, ASDF files are added to certain JWST calibration pipeline products or are part of the calibration pipeline reference data used by the software. In some cases, these provide with the distortion and spectral models needed to transform detector positions to a world coordinate frame along with the rules to combine them. Another use for these files is to provide with the necessary information to facilitate the creation of reference files.

For more information refer to the JWST Science Calibration Pipeline documentation.

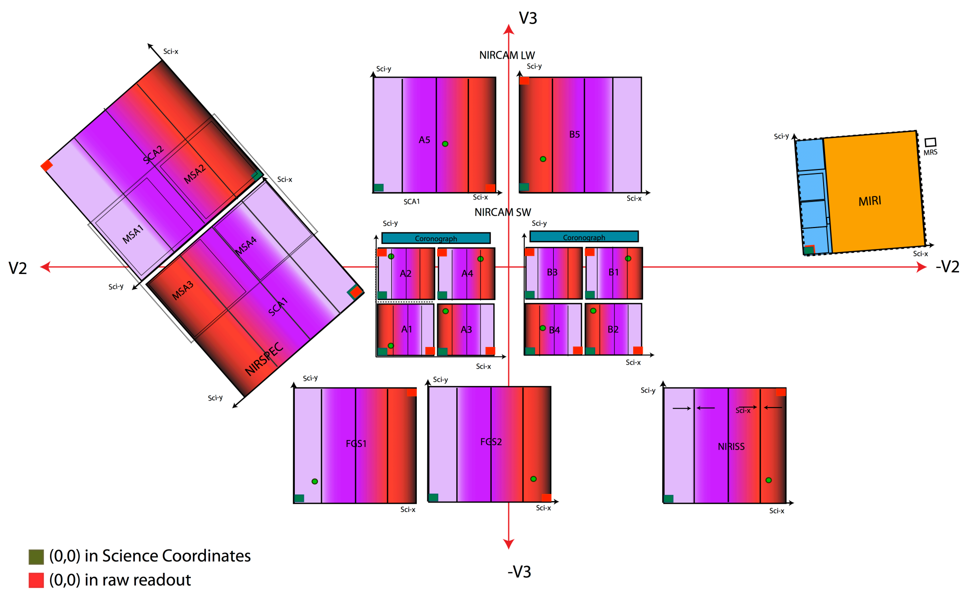
References

ASDF Standard 1.1.0
ASDF Astropy Documentation
JWST technical documents

Coordinate Systems and Transformations

The coordinates of JWST data are transformed from the raw coordinate system to the science coordinate system.
DMS produces data that has been rotated from the raw ("native" readout coordinate frame) orientations to the science orientations, putting all the detectors in the same orientation with respect to the V2, V3 axes in the focal plane. This requires flipping and rotating the reference pixels and science coordinates for NIRSpec, NIRCam, NIRISS, and FGS. MIRI by design has the same detector and science (X, Y) coordinates so no transformations are needed. The figure below shows with the red box the origin of the raw coordinates of the detector. The green box shows the origin of the science orientation, with the X and Y axis indicated with the black arrows. The lines across the the NIR detectors indicate the regions for each of the four output sublocks (512 × 2048) pixels in size, with the slow scan direction going opposite to the red box parallel to the black lines. The default fast scan directions are indicated with the arrows at the bottom of the NIRISS detector.

The orientation between the spacecraft V2/V3 coordinate system and the FITS science image x/y coordinate system is shown below. The telescope V1/V2/V3 coordinate system is right-handed with +V1 pointing toward the sky and out the page.



This figure compares the orientation of the raw images with those produced by DMS, also known as SCI orientation. The red box indicate the origin of the raw coordinates while the green box is the origin for the science coordinates. The black arrows are the axis for the science orientation. Although not shown here, the raw orientation is the same as the detector orientation, except for NIRISS. Note that NIRCam LW is offset. Science image frame orientations with respect to V2V3-frame V2 to the left and V3 up.

References

JWST-STScI-001256
JWST-STScI-002111
JWST technical documents

JWST Post-Pipeline Data Analysis

This article provides an overview of JWST post-pipeline data analysis tools and contains pointers to software installation and training materials.

Introduction

JWST post-pipeline data analysis tools are distributed as part of AstroConda to assist observers in viewing and analyzing their JWST data. The tools are generally written in Python and work with Astropy. Development is ongoing. All software is open source and community contributions are welcome in the form of suggestions, bug reports, or actual code. Further details on how to contribute can be found at the Data Analysis Tools Development Forum.

The suite of post-pipeline data analysis tools is intended to help astronomers with the often iterative and interactive workflow involved in converting these pipeline data products into meaningful scientific results. This involves tasks such as:

- inspecting data and data quality information;
- masking or flagging data and using those annotations to guide later steps in the analysis;
- using the results of interactive analysis to guide a custom run of the pipeline (e.g., tweaking spectral extraction parameters or background estimates);
- combining data sets in various ways, with careful attention to astrometry, PSF matching, and other issues;
- source detection and photometry using different choices or algorithms than those used in the pipeline;
- measuring lines and continuum in spectral data;
- fitting models to data or otherwise testing hypotheses.

A typical workflow involves highly interactive exploratory analysis on small portions of the data, followed by the development of custom scripts to automate the analysis on larger data sets.

Software installation

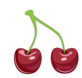
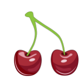
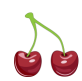
The recommended way to install stable versions of the JWST data analysis tools is to use AstroConda.

There are development versions of many of the tools linked in the "Repository" column in Table 1 that link to the open-source development locations on Github. Contributions to these developments are welcomed via bug reports (most effectively filed as github issues, but the JWST help desk is fine as well), or by code contributions through github pull requests. Use of the development versions of the code straight from github comes with the following caveats: at any given time, the code may not actually run or return correct results, and the documentation may be inconsistent with the code. Users who are not interested in contributing to the development software should use the versions in AstroConda.

Software packages

Table 1 provides links to further information about the tools. User training workshops are being offered to help new users familiarize themselves with these tools.

Table 1. Software tools information

Package	Purpose	Repository	Maturity
Astropy	A community python library for astronomy (documentation & tutorials)	Astropy	
glue	A Python library to explore relationships within and among related datasets.	glue	
ginga	A toolkit designed for building viewers for scientific image data in Python.	ginga	
Photutils	Tools for detecting and performing photometry of astronomical sources.	Photutils	
psfutils	Convenience tools for working with point-spread functions (PSFs)	psfutils	
astroimtools	Convenience tools for working with astronomical images.	astroimtools	
imexam	Tool for simple image examination, and plotting, with similar functionality to IRAF's imexamine.	imexam	

specviz	An interactive astronomical 1D spectra analysis tool with similar functionality to IRAF's splot.	specviz	
mosviz	A quick-look analysis and visualization tool for multi-object spectroscopy.	mosviz	
cubeviz	Interactive analysis tool for 3D spectroscopy (coming soon)	cubeviz	
asdf	Advanced Scientific Data Format is a next generation interchange format for scientific data (docs)	asdf	
gwcs	Generalized World Coordinate System tools for dealing with image and spectral geometries (docs)	gwcs	
synphot	Synthetic photometry toolkit for building model spectra and estimating count-rates. (docs)	synphot	

Levels of maturity run from prototypes with little or no documentation, symbolized by buds,

, progressing through various levels:

to

.

The exact meanings of these icons are a bit hard to quantify, but the flowers will tend to be still lacking in documentation and important features. The cherries are generally quite robust and well documented. The cherry pies are ready to be baked into your day-to-day workflow.

Be aware that all of the packages above are in very active stages of development, including Astropy and glue. For the ones at the cherry pie level, there is significant attention given to backwards compatibility as the APIs of the different modules evolve.

Training resources

There are many resources available for learning Python and for using Python for astronomical data analysis. This section provides pointers to astronomy-focused materials and to more JWST-specific training materials.

General python astronomical data analysis training

Table 2. List of training resources for python astronomical data analysis

Resource	Written materials	Videos	Comments
Using Python for Astronomical Data Analysis	Notebooks		January 2017 American Astronomical Society Special Session
Astropy	Tutorials		
Scientific Python Course at STScI	Notebooks	Videos	Notebooks are from 2015 version of course; videos from 2012-13
Astronomical Data Analysis with Python	Documents, notebooks	Videos	2014 Lecture series
Practical Python for Astronomers	Web documents		2011, 2012, 2013 Smithsonian Astrophysical Observatory

JWST focused training materials

Training materials for JWST data analysis will eventually include worked examples of common workflows, using outputs from the JWST pipeline. Currently the materials are more general than that, and make use example data sets from other observatories. Table 3 provides links to these materials.

Table 3. List of JWST focused training materials

Resource	Written materials	Videos	Comments
JWST User Training in Data Analysis	Notebooks	Webcast	2016 Workshop at STScI

Table 4 shows a more topic-oriented map of the materials from the 2016 Workshop.

Table 4. Introductory material

Introductory lectures		
Python	Notebooks	Webcast
Numpy	Notebooks	Webcast
Plotting	Notebooks	Webcast
Units	Notebooks	Webcast
Tables, I/O and FITS	Notebooks	Webcast
Coordinates	Notebooks	Webcast
Modeling	Notebooks	Webcast

Table 5. More specialized material

JWST user training		
JWST pipeline overview		Webcast starts at 50:00
Astropy & JWST tools overview		Webcast
Image examination (imexam)	Notebooks	Webcast
Image display toolkit (ginga)	Ginga embedded in a Jupiter Notebook	Webcast starts at 26:00 Glue-Ginga Interoperability Webcast 1:03:00
Photometry (photutils)	Intro Notebooks PSF-photometry Notebook	Webcast starts at 55:30 PSF-fitting Webcast 46:00
Data exploration (glue)		Webcast Customizing Glue Webcast 27:30
Interactive spectral analysis (SpecViz, MosViz, CubeViz)		Webcast starts at 1:01:30
NIRCam Data Reduction		Webcast
Advanced Astropy Tables	Notebook	Webcast
Generalized World Coordinate System (gwcs)	Notebook	Webcast

Obtaining JWST Data

All JWST data products are archived in the Mikulski Archive for Space Telescopes (MAST) and can be searched and explored via the MAST Discovery Portal. Authorized users may visualize and retrieve queried data.

Introduction

Data from all JWST instruments: MIRI, NIRCam, NIRSpec, NIRSS, and FGS, are archived in the Mikulski Archive for Space Telescopes (MAST). The data products include raw exposures and all levels of calibration processing, contemporaneous engineering and calibration reference data, and in some cases high-level science products contributed by investigating teams. All JWST data products can be searched via the MAST Discovery Portal, or through various community applications and software interfaces. Authorized users may visualize and retrieve queried data.

The JWST data archive: MAST

Data from all JWST instruments are archived in MAST, which is the permanent repository and access portal for data from many NASA UV/optical/IR missions and selected ground-based surveys. Users may discover particular JWST data in a variety of ways, including searches by:

- the observing program in which they were obtained
- the instrument used to obtain the data
- sky coordinates
- astrophysical source name

MAST Discovery Portal users can find all varieties of JWST data products along with co-spatial products from any other hosted mission archive. Authorized users may explore the JWST science data products, overlay science datasets from other missions and surveys, as well as browse the contemporaneous engineering data, calibration data, and data quality summaries. Authorized users may also retrieve JWST data products via the MAST web interface or through user-written scripts by calling the MAST Application Programming Interface (API).

The following sections describe the means to discover, explore, and retrieve JWST data products, and to obtain user support.

Data Discovery

The MAST Discovery Portal is the primary web interface for discovering, visualizing, assessing, and retrieving calibrated data products from the JWST calibration pipeline, as well as calibration and engineering data products and services.

Introduction

There are a variety of ways for researchers to access JWST data of interest, including the MAST Discovery Portal which is the primary web interface for discovering, visualizing, assessing, and retrieving archived data. The portal also provides access to ancillary and engineering data related to the observations, and data quality summaries that enable users to assess any anomalies during the observations, the achieved exposure depth, and the calibration processing. Users may evaluate contemporaneous calibration reference data that were used to remove the instrumental signature from the science data products. They may also discover data through a programmatic (software-based) interface such as the MAST Applications Programming Interface (API), and through various community tools. Users may elect to be notified as data from new JWST observations become available through MAST.

Data product types

The JWST Data Management System (DMS) produces many products for each JWST observation, including the science files produced by the data reduction pipeline. The exact type and number of products depends upon the instrument, its configuration, and operating mode. Consult the JWST File Names, Formats, and Data Structures article for a detailed description of each science product and the concomitant data. Most of the science data files are images or tables in FITS format (Pence, et al. 2010), while others are in structured or unstructured ASCII format. Table 1 contains a short summary of the data product types that may be included with each dataset. Table 1 below summarizes the semantic content of the various data products, including some that are produced outside the science data reduction pipeline.

Table 1. Summary of data product types that may be included with each dataset

Product type	Format	Description
Science data	FITS MEF images	All levels of science data products from imaging, spectroscopy, and time-series observations. The image extensions contain multi-dimensional science pixel arrays, concomitant data quality flags, and variance arrays.
	FITS BINTABLEs	Extracted spectra are stored as tables, and include fields for spectral coordinates, flux, and concomitant quality information.
	FITS FOREIGN	Metadata for images or spectra are stored in ASDF format, and are packaged in a FITS extension of type FOREIGN.
Source catalog	FITS binary table	Measurements of detected source properties, derived from image exposures.
Associations	JSON	Enumeration of relationships between all levels of science data products, and related calibration reference products, that were used during the course of calibration pipeline processing.
Association summary report	HTML	A summary of associated exposure and observation attributes, graphics, plots, and completion status.
OPD	FITS	Optical Path Difference images that describe the state of the mirror segment alignments for a particular epoch. The OPD files can be used with a tool such as WebbPsf to determine the estimated PSFs for any given instrument configuration.

Minimum recommended data products

Of the many different data products produced by the calibration pipeline, a subset has been identified as essential for extracting the intended science from the data. These are termed the "minimum recommended data products" (MRDP). The selection of data products that are included in this set depends upon the instrument used to obtain the data, and its configuration and operating mode. Generally, products in the MRDP include the lowest-level science product for which the instrumental signature has been removed along with the associated data quality products, but exclude calibration and certain ancillary products.

> If close examination of the MRDP or the data quality flags in the report indicates problems, it may be wise to retrieve the entire dataset for review and (possibly) custom processing.

Observing program assessment

Observing programs for JWST can be highly complex, consisting of multiple exposures, or observations of multiple targets or fields, some of which are related in the data processing through associations to produce combined or stacked products. The exposures may consist of, e.g.:

- Dithered exposures
- Tiled imaging over an area of sky, in one or more bands
- Coronographic imaging, including a PSF reference star

Details of the exposures, their relationship to one another, their context in the overall observing program, and the status of their completion for active programs) are summarized in the JWST Associations Summary Report.

Data reprocessing

Data from JWST are likely to be reprocessed with the calibration pipeline from time to time, particularly early in the mission lifetime. Users may also wish to reprocess JWST data for many of the same reasons, including:

- Additional observations of the target/field become available
- Better or more appropriate contemporaneous calibration reference files are created
- New software or calibrations are implemented to remove unexpected instrumental signatures
- Custom processing is required to achieve the optimal scientific result

The provenance metadata in the file headers will indicate the date, software version, and (unique) calibration file identifiers. In addition, release notes for the pipeline software are published, so that the epoch of processing by STScI will be apparent.

MAST Discovery Portal

MAST implements various protocols of the Virtual Observatory (VO) including those for image, table, and spectral data access, and the Common Archive Observation Model (CAOM) that describes the core services offered by MAST. These are the mechanisms by which MAST data can be searched and retrieved by VO-aware applications, and conversely, the mechanisms for the Discovery Portal application to discover, retrieve, and operate on data not hosted in MAST.

Archived data

A portal has been developed for MAST data collections, including JWST. The MAST Discovery Portal offers great flexibility in customizing queries to identify datasets to explore and retrieve. While the interface is straightforward, a user guide may help first-time users. See the Data Exploration article for a detailed example of a search using the Discovery Portal.

Anticipated data

At any given time some observations in a program may have been executed, archived, and become available to the community; some archived observations may temporarily be restricted to those with exclusive access; while still other observations may remain to be obtained. Investigating teams and the broader community each have an interest in data availability. In order to encourage the greatest possible use of JWST data in MAST, a subscription service will notify registered and subscribed users when one of the following observation-related events occur:

- new observations are Archived,
- new, relevant calibration reference files become available,
- archived data have been reprocessed, or
- restricted-access data become available to the public.

Users may tune the notifications by mission, program ID, event type, and science product type. Users may establish or cancel their subscription through the MAST Discovery Portal, change the media and frequency of notifications, and change the selection criteria for notifications.

Duplication checking

MAST provides the capability to check a list of potential target observations against the extant and planned observations to check for duplications. Duplicated observations are, in general, not allowed so it is in the proposer's best interest to perform this check prior to submitting a proposal for review.

Engineering data

Engineering database service

The engineering database service provides an archival repository of calibrated engineering telemetry from the instruments and spacecraft. These data are much more detailed than the subset that is represented in the headers of science FITS files. This interface also provides a view into the status of the science data processing and calibration pipelines. The JWST Archive interface provides user access to this database.

Virtual observatory tools

Many community software applications are capable of accessing remote data using protocols developed for the Virtual Observatory. You may have used them before without being fully aware of how such data were obtained. Table 2 provides an incomplete list of community, VO-aware applications for visualizing and exploring archived astronomical images, spectra, and catalogs:

Table 2. An incomplete list of community, VO-aware applications for visualizing and exploring archived astronomical data

Application	Type	Description
Aladin	Analysis	Interactive sky atlas, capable of overlaying archived images and source catalogs
DataScope	Discovery	Position-based search engine for astronomical archives
SAOimage/DS9	Analysis	Image display tool, capable of overlaying archived source catalogs
SIMBAD	Retrieval	Astronomical catalog search engine
SkyView	Discovery	Astronomical image search engine
SpecView	Analysis	1-D spectrum visualization and analysis
TOPCAT	Analysis	Catalog cross-match, filtering, sub-setting, and visualization tool
ViszieR	Retrieval	Astronomical catalog search engine
astropy.vo.client	Discovery	Python packages for using VO services, including accessing VO registries, as well as catalog manipulation in software

References

Pence, W. D., Chiappetti, L., Page, et al. 2010, A&A, 524, A42
Definition of the Flexible Image Transport System (FITS), version 3.0

JWST technical documents

Data Exploration

Introduction

Investigators have a variety of tools at their disposal with which to explore JWST data, and data from other missions and surveys, both within and outside of the context of the MAST Discovery Portal. Web- or desktop-based tools such as those for viewing areas of digital sky surveys, selecting images and footprints from popular missions and surveys, cross-matching and plotting source catalogs, and viewing timeseries of selected targets enable Archive users to plan JWST observing programs, and to place JWST new observations in the context of the extant astronomical data from around the world.

Tools for data exploration

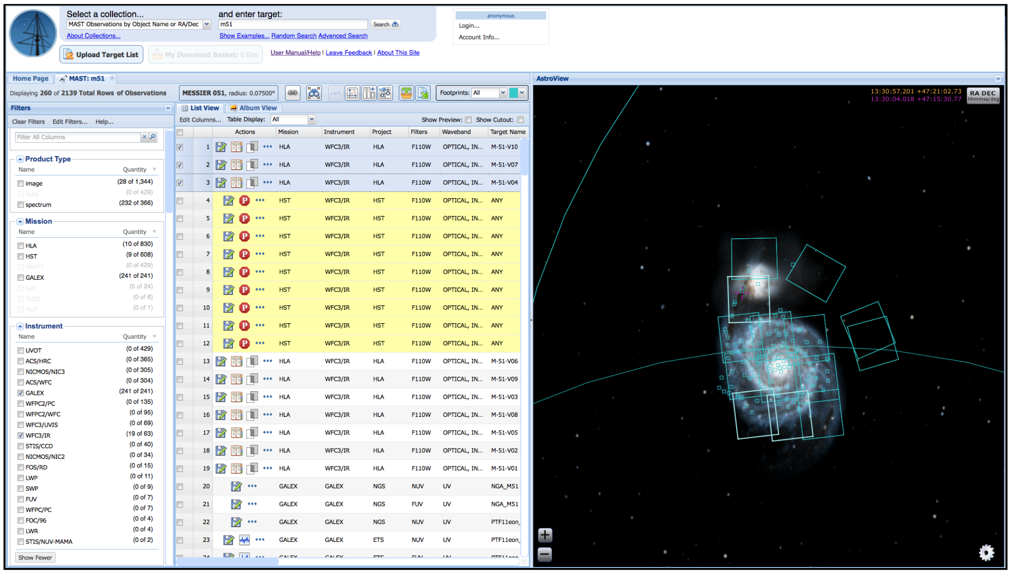
Once a search for data has been performed, there are a number of tools available to help sort, visualize, and refine the results. Many of these are built into the MAST Discovery Portal (DP), but there are also many additional tools available from the community at large. These tools enable users to plan JWST observing programs, explore JWST data once the observations are available in the Archive, and understand JWST data in the context of the world-wide collection of data on a given source or field. Below are descriptions or links to tutorials on how to use these tools.

Native MAST tools

Target/Position query

A simple query of a target name or position in the MAST DP will often bring up an abundance of information, with many ways to sort and filter in order to get exactly the data you are looking for.

Figure 1. Result of a Data Discovery Portal query around M51, filtered for GALEX and WFC3/IR images

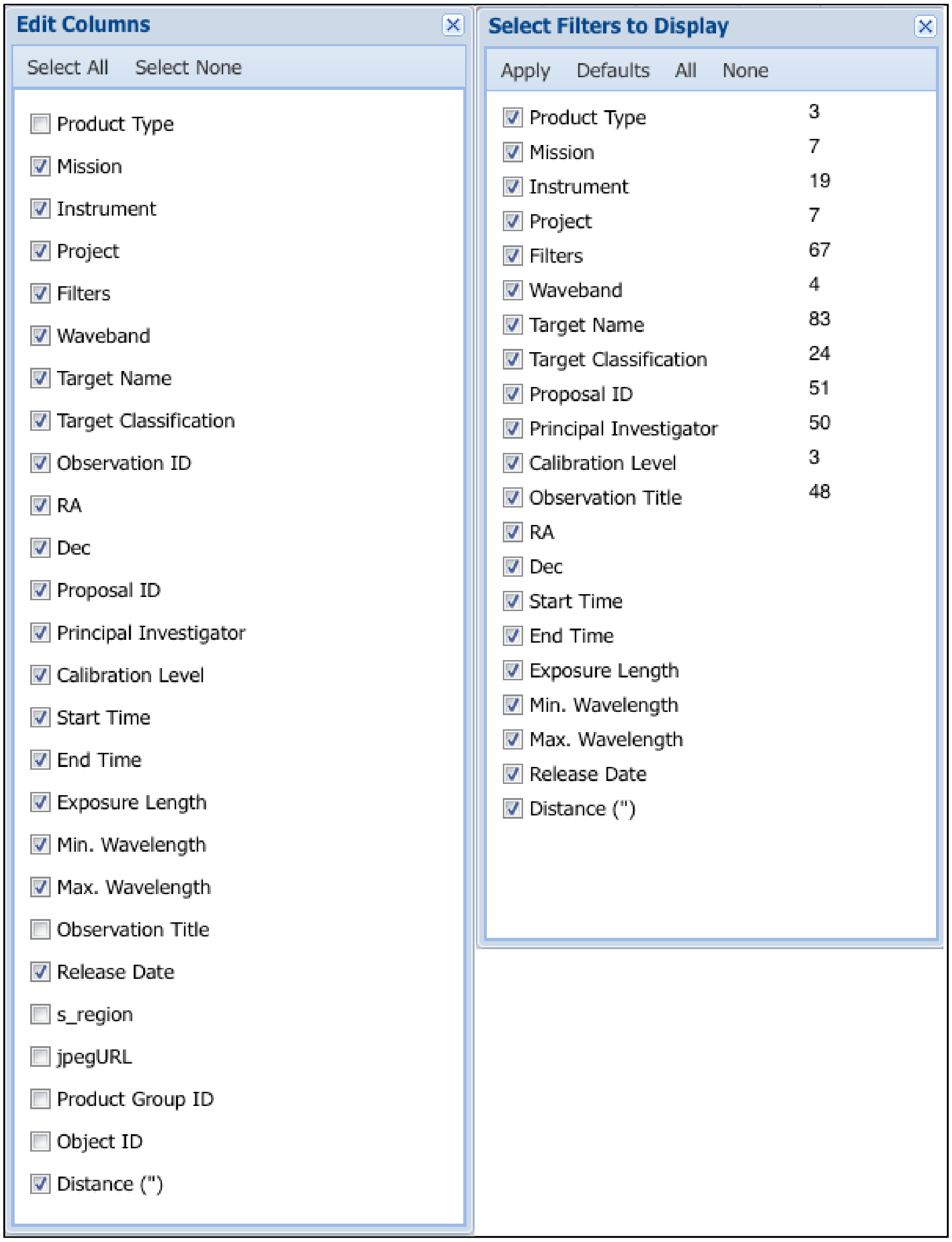


The top of each search results tab provides a number of tools for the user, including buttons for adding selected files to the download basket and exporting the search results table, as discussed in the Data Retrieval section. The user is also given the ability to cross-match the current search results with another catalog from either MAST or CDS, and to create custom plots from various parameters within the current search results.

The leftmost panel of the results page provides a number of filters to isolate specific data characteristics of interest. These include product types (image, spectrum, etc.), mission (HST, SWIFT, FUSE, etc.), instrument, filters, waveband (optical, UV, infrared, etc.), principal investigator, calibration level, exposure length, start / end time, and several others. These filters can accommodate a wide range of specific requests and combinations of constraints for narrowing down a large set of search results.

The List View in the middle of the results page provides a search results table with detailed information for all current search results after the application of any filters the user has selected. Each column in the table may be used to sort the results as the user wishes, and the columns themselves may be customized to suit the user's needs (Edit Columns button in the upper-left corner of the panel). Image previews may also be added to the table with the Show Preview checkbox. For larger views of these preview images, the user may also switch from the List View tab to the Album View tab, where the same search results are laid out in a tiled array. User sorting may not be modified in this view, but it will maintain any sorting parameters the user has implemented in the List View table.

Figure 2. All available List View columns (left) and Filter options (right)



The right side of the MAST Portal search results displays a built-in AstroView panel, which is updated to display the STScI Digitial Sky Survey (DSS) imaging of the target area overlaid with footprints of the search result images. This panel is active and may be zoomed and panned as the user wishes, and will update in real time as the user filters out various search results. The user may even click on individual footprints in this panel to select specific images in both the list and album views of the search results. The AstroView panel also provides some settings options for customizing aspects of the display, such as display of the given search radius, and RA/Dec grid lines.

DP Catalog Viewer

The DP Catalog Viewer tool enables visualizing source catalogs, including the Hubble Source Catalog (HSC; Whitmore et al. 2016), as positions distributed on a displayed image or as a user-defined plot of various source attributes. A number of tutorials have been developed for various science uses, and some are highly relevant for JWST data users.

Tutorial	Description
Catalog cross-match	Perform cross-match between user catalog and the HSC
Catalog Filtering	Isolate the red sequence in the galaxy cluster Abell 2390
Catalog filtering and outlier detection	Search for variable objects in the dwarf irregular galaxy IC 1613
Outlier detection with the HSC	Search for white dwarfs in the globular cluster M4 using the HSC and the CasJobs engine

JWST Engineering Database Browser

TBA: discussion of the interface for the JWST engineering database once it is developed.

Community data visualization tools

Below is an incomplete list of community, VO-aware tools for visualizing and exploring archived astronomical images, spectra, and catalogs.

Application	Type	Description
Aladin	Analysis	Interactive sky atlas, capable of overlaying archived images and source catalogs
DataScope	Discovery	Position-based search engine for astronomical archives
SAOimage/DS9	Analysis	Image display tool, capable of overlaying archived source catalogs
SIMBAD	Retrieval	Astronomical catalog search engine
SkyView	Discovery	Astronomical image search engine
SpecView	Analysis	1-D spectrum visualization and analysis
TOPCAT	Analysis	Catalog cross-match, filtering, sub-setting, and visualization tool
ViszieR	Retrieval	Astronomical catalog search engine
astropy.vo.client	Discovery	Python packages for using VO services, including accessing VO registries, as well as catalog manipulation in software

References

Fleming, S. W., et al. 2015, AAS, 225, 336.59
Beyond the Prime Directive: The MAST Discovery Portal and High Level Science Products

Whitmore, B., et al. 2016, AJ, 151, 134
Version 1 of the Hubble Source Catalog

JWST technical documents

Data Retrieval

Introduction

Data identified with a search may be retrieved in a number of ways, including the MAST Portal download manager, HTML download, ftp, and with user software. The best method of retrieval is largely a matter of convenience, although requests for very large volumes of data may require special effort. Only authenticated and authorized users may retrieve data within the exclusive access period.

Retrieving JWST data from MAST

Data that have been identified through a search may be retrieved in a number of ways by authorized users. The best method of retrieval is largely a matter of convenience, although requests for very large volumes of data may require special effort.

Anonymous users may access public datasets in MAST without registering and without logging in. But only authenticated and authorized users may access data within the exclusive access period. See the Data Access Policy for details and information on data access restrictions.

MAST portal retrieval options

Users who have identified data of interest within the MAST Data Discovery Portal may retrieve the files (or collections of files) by placing them in the "basket" and clicking the "save" icon. See Portal UI Data Retrieval for details. The MAST Portal also allows for exporting of results data tables and some batch retrieval options including DVD shipment, or staging data for later download. The standard download basket procedure offers the user a choice of zip, tar, tar.gz, or cURL options for the download package, but stipulates a 5.12 GB total file volume maximum. Above this limit, the user will be obligated to use the cURL script option or the MAST API within a script (see below) for a non-batch retrieval. Batch retrieval options may be ideal if exceeding this volume limit, though this may be true of smaller sizes depending on the user's network connection.

Table 1. Batch retrieval methods

sFTP	• The user must provide a server hostname and directory, along with a server username and password to facilitate data transfer. • The user must ensure there is adequate disk space available on the destination host and directory they have designated.
staging	• Data will be staged to an STScI FTP server. • The user will be notified when the transfer is complete and will be given the server location.
DVD	• The user must provide a full mailing address for shipment. • Overnight shipping (where available) once the disk is prepared. • DVDs should in most cases be delivered within a week of the request • Likely only useful for users with very limited network access • Note: The maximum capacity of a DVD (~5 GB) may be too small for the largest JWST raw exposures.

Note that a valid email address must be provided for all batch data retrieval requests, though an STScI MyST ID is not required for public data.

MAST portal download examples

Figure 1. A MAST Portal query of M51 for HST data

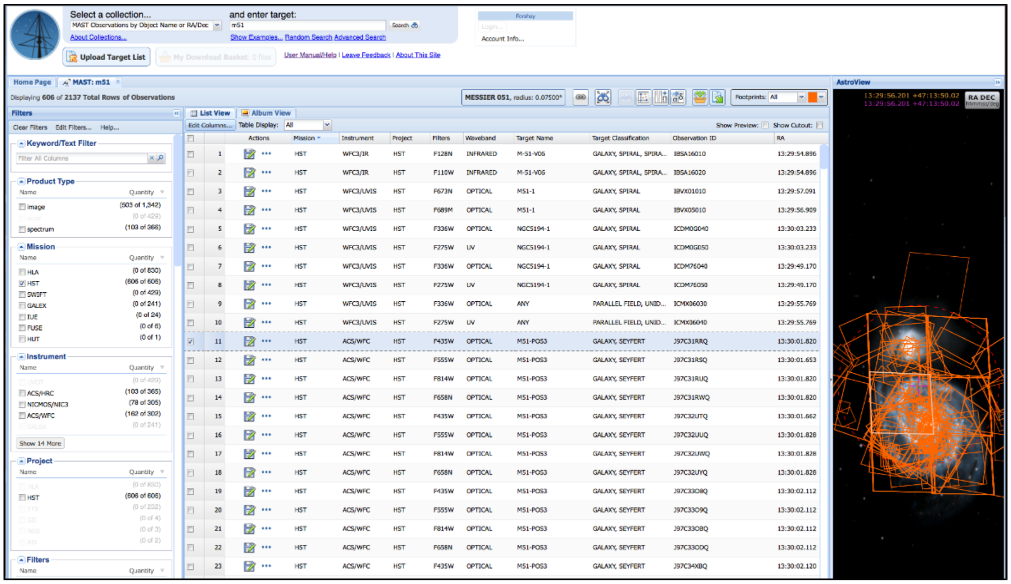


Figure 2. Selected images are added to the download basket with Minimum Recommended Products still enabled

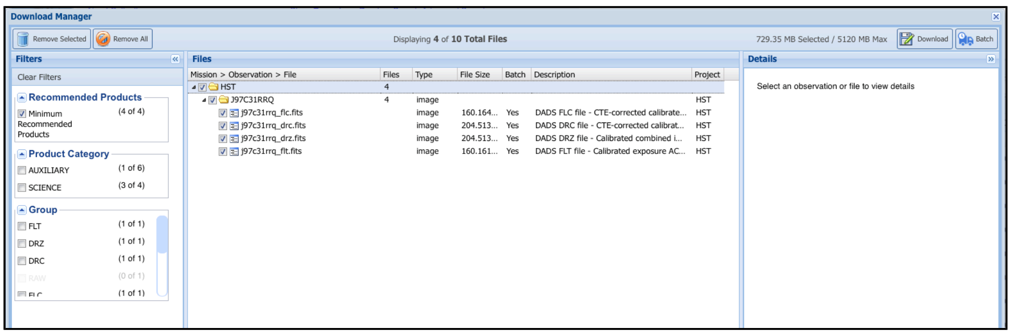


Figure 3. Download options for the selected images

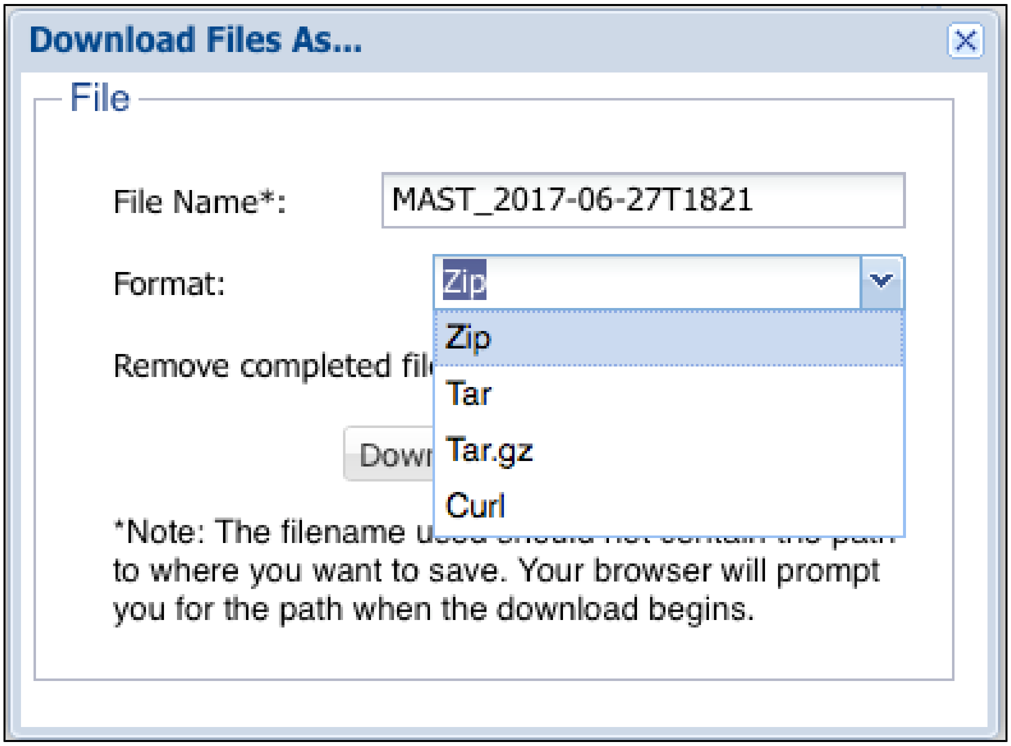


Figure 4. Batch retrieval options for selected images

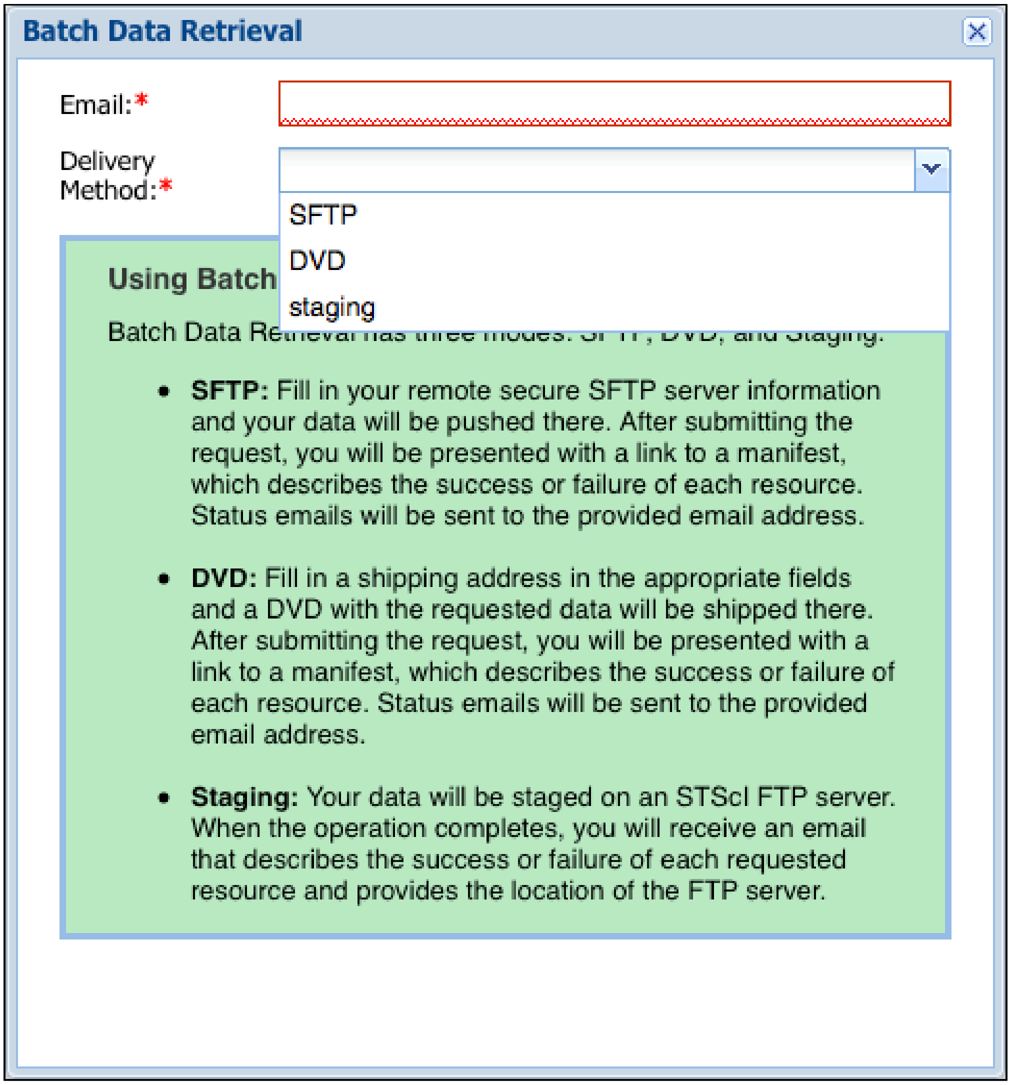
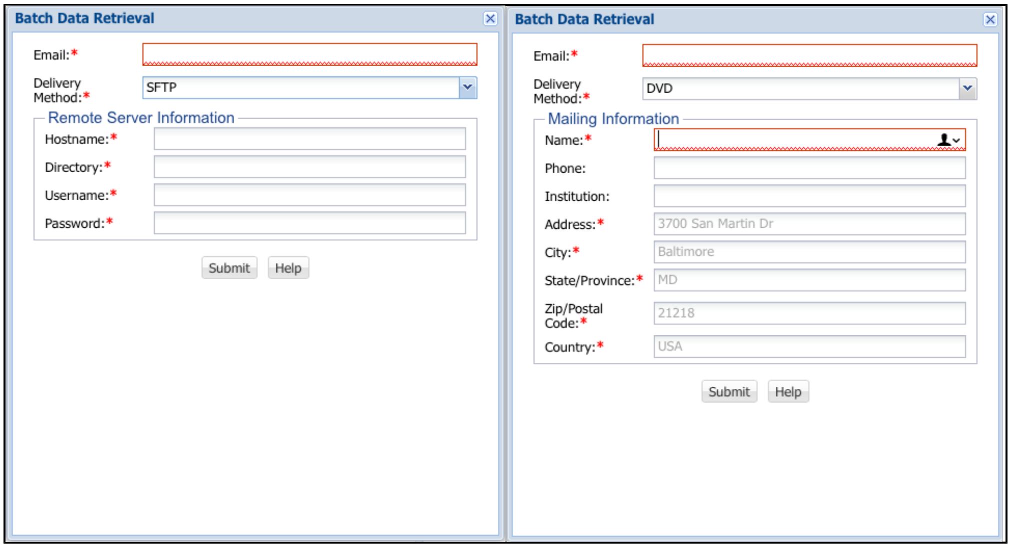


Figure 5. Required fields for sFTP retrieval and physical media delivery



Note the username & password in the sFTP retrieval are for the destination server, NOT your STScI MyST credentials.

The MAST programmatic API

It is possible to retrieve data using scripts in various programming languages, such as Unix shell scripts or Python, through the Mashup API for MAST. See Programmatic Interfaces for details.

Community tools

A variety of community applications can retrieve data from astronomical archives, including MAST. These tools use Virtual Observatory protocols to discover and retrieve data files. Table 2. proviedes with a list of some of them.

Table 2. List of applications and cabilities

Tool	Capabilities
Aladin	Sky atlas tool
SAOImage/DS9	Image and catalog display tool
TOPCAT	Catalog joins, subsetting, editing, and visualization
VizieR	Extensive VO catalog offerings
VOClient	Standalone software for querying and downloading VO data from USVAO

> These tools currently lack the ability to provide user authentication credentials to archive services, and so cannot be used to retrieve data archived within the exclusive access period.

VizieR retrieval examples

Figure 6. Searching the hst/obscore catalogue in VizieR for the same obsID as the MAST Portal example

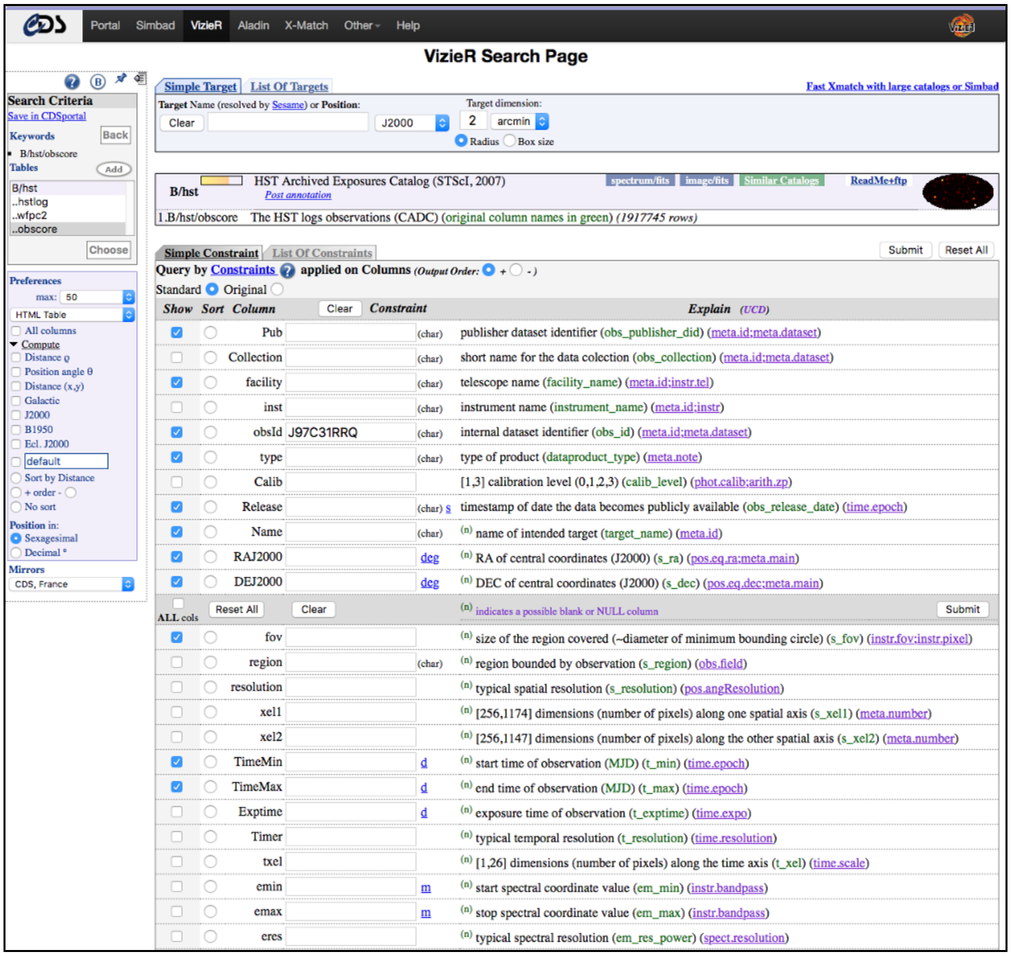


Figure 7. VizieR is able to find the same M51 data found through MAST Portal

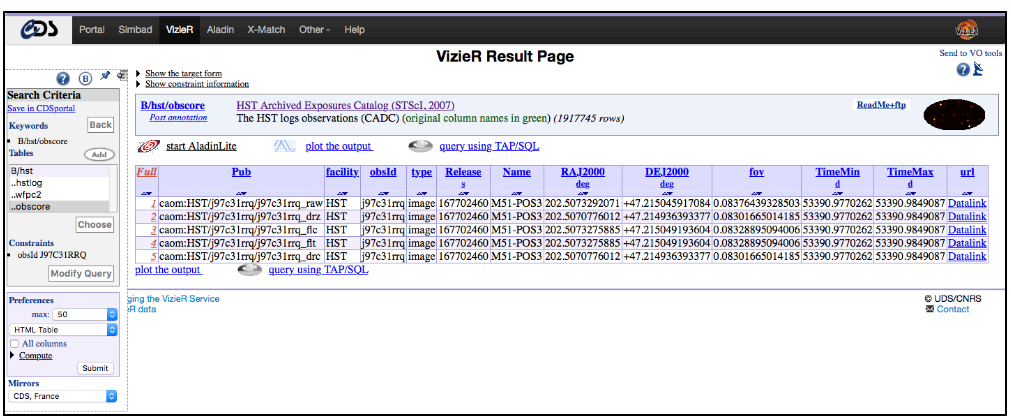


Figure 8. The resulting VOTable downloaded for the _drz file


```xml
<?xml version="1.0" encoding="UTF-8"?>
<VOTABLE xmlns="http://www.ivoa.net/xml/VOTable/v1.3" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" version="1.3">
  <RESOURCE type="results">
    <TABLE>
      <FIELD name="ID" datatype="char" ucd="meta.id;meta.main" arraysize="*" />
      <FIELD name="access_url" datatype="char" ucd="meta.ref.url" arraysize="*" />
      <FIELD name="service_def" datatype="char" ucd="meta.ref" arraysize="*" />
      <FIELD name="error_message" datatype="char" ucd="meta.code.error" arraysize="*" />
      <FIELD name="semantics" datatype="char" ucd="meta.code" arraysize="*" />
      <FIELD name="description" datatype="char" ucd="meta.note" arraysize="*" />
      <FIELD name="content_type" datatype="char" ucd="meta.code.mime" arraysize="*" />
      <FIELD name="content_length" datatype="long" ucd="phys.size;meta.file" unit="byte" />
      <FIELD name="readable" datatype="boolean">
        <DESCRIPTION>the caller is allowed to use this link with the current authentication</DESCRIPTION>
      </FIELD>
      <DATA>
        <TABLEDATA>
          <TR>
            <TD>caom:HST/j97c31rrq/j97c31rrq_drz</TD>
            <TD>http://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/data/pub/HSTCA/j97c31rrq_drz?runid=ad5nqzq5sre8axlw</TD>
            <TD />
            <TD />
            <TD>#this</TD>
            <TD />
            <TD>application/fits</TD>
            <TD>214496640</TD>
            <TD>true</TD>
          </TR>
          <TR>
            <TD>caom:HST/j97c31rrq/j97c31rrq_drz</TD>
            <TD />
            <TD>soda-f4ed9ec4-6695-43d2-9ee5-0d0296650274</TD>
            <TD />
            <TD>#cutout</TD>
            <TD />
            <TD>application/fits</TD>
            <TD />
            <TD>true</TD>
          </TR>
          <TR>
            <TD>caom:HST/j97c31rrq/j97c31rrq_drz</TD>
            <TD />
            <TD>soda-3c20e971-670f-47da-86d8-bc2853209434</TD>
            <TD />
            <TD>#cutout</TD>
            <TD />
            <TD>application/fits</TD>
            <TD />
            <TD>true</TD>
          </TR>
          <TR>
            <TD>caom:HST/j97c31rrq/j97c31rrq_drz</TD>
            <TD>http://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/data/pub/HSTCA/j97c31rrq_prev.fits?runid=ad5nqzq5sre8axlw</TD>
            <TD />
            <TD />
            <TD>#preview</TD>
            <TD />
            <TD>application/fits</TD>
            <TD>14226447</TD>
            <TD>true</TD>
          </TR>
          <TR>
            <TD>caom:HST/j97c31rrq/j97c31rrq_drz</TD>
            <TD>http://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/data/pub/HSTCA/j97c31rrq_prev.jpg?runid=ad5nqzq5sre8axlw</TD>
            <TD />
            <TD />
            <TD>#preview</TD>
            <TD />
            <TD>image/jpeg</TD>
            <TD>803220</TD>
            <TD>true</TD>
          </TR>
          <TR>
            <TD>caom:HST/j97c31rrq/j97c31rrq_drz</TD>
            <TD>http://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/data/pub/HSTCA/j97c31rrq_flt_hlet?runid=ad5nqzq5sre8axlw</TD>
            <TD />
            <TD />
            <TD>#auxiliary</TD>
            <TD />
            <TD>application/fits</TD>
            <TD>50705</TD>
            <TD>true</TD>
          </TR>
          <TR>
            <TD>caom:HST/j97c31rrq/j97c31rrq_drz</TD>
            <TD>http://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/data/pub/HSTCA/j97c31rrj_jif?runid=ad5nqzq5sre8axlw</TD>
            <TD />
            <TD />
            <TD>#auxiliary</TD>
            <TD />
            <TD>application/fits</TD>
            <TD>3661</TD>
            <TD>true</TD>
          </TR>
          <TR>
```

Programmatic Interfaces

JWST data may be queried and accessed via community tools or user-written programs through an Application Programming Interface, such as the MAST API, or various Virtual Observatory tools.

Introduction

Programmatic interfaces allow a user to discover and access JWST data through scripted queries in various programming languages, instead of going through the standardized web interface. This affords a greater degree of complexity and customization in searches and retrievals. Through this mechanism the user may also cast a wider search net in an effort to discover any and all data related to a particular object or location. The two primary methods of performing these queries are through the MAST API or various Virtual Observatory services.

MAST API

The Mashup Application Programming Interface (API) provides a way to access data stored in the Mikulski Archive for Space Telescopes (MAST) through a programming language of the user's choice. It does this by translating customized URL requests into database queries, which allows the user to create scripts designed to assemble specific query parameters and to send those queries in batches.

Beyond simple queries, the MAST API is able to extend much of the functionality of the MAST web portal to a programming environment. This includes the ability to perform cross-matching searches with various data catalogs, cone searches of a specific target location, table filtering based on specific columns, and target name resolution. For detailed service descriptions, instructions, and examples of this API in use, please refer to the Mashup API page. This page also includes a Jupyter notebook for a customizable example of a Python MAST data search and retrieval.

VO Services

The NASA Astronomical Virtual Observatories (NAVO) program aims to provide access to data stored in some of NASA's largest archives by creating a standardized interface for submitting queries across multiple datasets. These archives include MAST along with the High-Energy Astrophysics Science Archive Research Center (HEASARC), the Infrared Science Archive (IRSA), and the NASA Extragalactic Database (NED).

Similar to the Mashup API, VO services also primarily function by translating customized URL requests into database queries. The four primary types of available queries follow:

- Simple Cone Search (SCS) – Allows a user to search for observations around a specified set of coordinates within a given radius.

- Simple Image Access Protocol (SIAP) – Provides query access to 2-dimensional images and 3-dimensional or greater data cubes.
- Simple Spectral Access Protocol (SSAP) – Provides query access to 1-dimensional spectral data.
- Table Access Protocol (TAP) – Allows a user to perform complex queries on data tables in the Astronomical Data Query Language.

SCS, SIAP, and SSAP queries all provide broad methods of data discovery within certain search parameters, while TAP is better suited to finding specific data with more tightly-constrained queries.

More information on VO access to MAST holdings can be found here. Table 1 lists additional VO resources and tools.

Table 1. Additional VO resources

AstroGrid	VO services hosted at the Institute for Astronomy, University of Edinburgh. (last updated 2012, defunct?)
Canadian Virtual Observatory	Various VO tools hosted by the CADC.
CasJobs	A tool for creating custom databases using MAST and GALEX data.
DataScope	Target name and position searches hosted by HEASARC.
International Virtual Observatory Alliance	The IVOA sets standards and facilitates international VO cooperation. A good deal of VO information and links to more tools can be found here.
seleste	GUI and command line VO tools from the Smithsonian Astrophysical Observatory
Virtual Observatory Services	A collection of VO tools developed by researchers at Johns Hopkins University.
VOClient	Standalone VO software from the USVAO.

Data Access Policy

Introduction

Science data obtained from JWST will be released to the astronomical community following an exclusive access period, during which the Principal Investigating Team enjoys exclusive scientific use. The duration of the exclusive access period depends upon the program. Access to data during this period is restricted to authenticated and authorized users. Principal Investigator team members and MAST users should acknowledge the source of the data and the funding grant (if any) in any publication using JWST data.

More about the data access policy

Access to and use of JWST science data is governed by a mission policy not unlike that adopted for HST. Science data obtained from JWST will be released to the astronomical community and the public following an exclusive access period. Access to archived data during this period is limited to authenticated and authorized investigators. Principal Investigator team members and MAST users should acknowledge the source of the data and the funding grant (if any) in any publication using JWST data.

Exclusive access period

JWST science data become available to the astronomical community and the public upon expiration of an exclusive access period, during which the investigator team enjoys exclusive access to these data in MAST. This period applies at the data product level (rather than to a completed observing program), and begins the moment that individual exposures obtained from the spacecraft have been calibrated and made available in MAST. The duration of the exclusive access period depends in part upon the type of observing program used to obtain the data, as summarized in the following table.

Table 1. Exclusive access period by program type

Program Type	Exclusive Access Period
Early-Release Science	None
GTO/GO Science	12 months
Calibrations	None

Embargoed Data
Data obtained during JWST science commissioning, possibly including ERS program exposures, may be embargoed for a limited period of time following the observation to verify the performance of the system and the data processing, or for the preparation of public outreach materials.

Principal Investigators (PIs) have the option of waiving all or part of their exclusive access period; in certain limited cases the STScI Director may grant a request for an extension. Note that all data, whether or not access is restricted within MAST, may be used by authorized JWST project staff for monitoring and diagnostic purposes. See the policy on Data Rights and Data Dissemination for details.

Access to exclusive access data in MAST requires users to establish electronic credentials with STScI; browsing and retrieval of such data requires both authentication of the user credentials and authorization to access data from particular programs.

Higher-Level Data Products

The period of exclusive access for some higher-level data products is a little more complicated. This is because some data products may be created by combining data from multiple observations that are obtained over a (possibly extended) period of time. The data that contribute to a combined product are associated, and if these associations change because new observations become contributors, the higher-level product is automatically recreated when the new observations are processed. In this case the exclusive access period is determined by the most current dataset that contributes to the associated high-level product. For more information about these cases, please refer to the associations documentation.

This means that, although some calibrated exposures that contribute to an association may be public, the higher-level product of the association will not be public if the most recent exposures within the association fall within the exclusive access period. This creates the somewhat odd possibility that certain high-level products (e.g., a stage-3 tiled image mosaic of an extended field) might revert to exclusive access, even if they were previously public. Note that stage 2c data products (see Data Product Naming Conventions) represent a special case. These calibrated exposure-level products contain data (actually, concomitant data quality values) that originate from all the science exposures that are included in the association. As such the period of exclusive access for these products will be governed by that for the most recent observation from which they were created.

Archive registration

Investigator team members who wish to access data within the exclusive access period must establish a STScI MyST account to establish electronic credentials to authenticate their identity. PIs and CoIs of JWST observing programs already have MyST credentials, which they used to submit their proposals for review. Investigator team members must also be authorized to access data from their programs. PIs will automatically receive such authorization, which they may also grant to any collaborator who is registered with the STScI MyST.

> Data for which the exclusive access period has expired may be retrieved anonymously from MAST. A MyST account is not required to search, view, explore with MAST tools, or retrieve data that have been released to the public.

Acknowledgements

It is important for investigators who publish scientific results based on MAST data to acknowledge the mission and (if applicable) the source of funding. The appropriate acknowledgement depends upon which data are used. All publications based on JWST data must include the following footnote (generally somewhere near the beginning of the paper):

> Footnote for publications based on JWST data:
> Based on observations made with the NASA/ESA James Webb Space Telescope, obtained from the data archive at the Space Telescope Science Institute. STScI is operated by the Association of Universities for Research in Astronomy, Inc. under NASA contract NAS 5–26555.

In addition, if your JWST archival research was supported by a grant from STScI the publication must also carry the following acknowledgment, including the grant number, at the end of the text:

> Acknowledgement for Archival Research using JWST data:
> Support for this work was provided by NASA through grant number ________ from the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS 5–26555.

Publications based on all other data (i.e., not from JWST or HST) obtained from MAST should include the following acknowledgement:

> Acknowledgement for research using other data from MAST
> Some/all of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute. STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5–26555. Support to MAST for these other data is provided by the NASA Office of Space Science via grant NAG5–7584 and by other grants and contracts.

Publication statistics are important to the JWST mission. Please send a link to each preprint (e.g., in arXiv), or final refereed publication, that is based on JWST data to the following address:

> Library address for STScI
> Librarian
> Space Telescope Science Institute
> 3700 San Martin Drive
> Baltimore, MD 21218
> library@stsci.edu

Archive User Support

Introduction

MAST staff members assist science users of JWST data (and data from other hosted missions) with searches, notifications, the use of visualization tools, and standard or scripted data access and retrievals. Archive Contact Scientists will in addition assist JWST Observer teams and Archive Researchers with the preparation of high level science data products to be hosted at MAST.

User support for MAST

MAST archive staff members assist the general science community with searches for data, visualization with Portal components, data retrievals, and also assist Principal Investigators with establishing authorization for their team members to access to their proprietary data. Archive Contact Scientists assist JWST PI and Archival Research teams in preparing High-Level Science Products (HLSPs) that will be hosted by MAST.

High level science products

HLSPs are required deliverables of funded JWST Archival Research and Early Release Science programs (see the ERS proposal preparation instructions), but are encouraged from any science team that uses JWST data as a part of their research. These data products should provide substantial science value to the general astronomical community, and have broader appeal beyond the original scope of the science goals for a particular program, or should provide technical insight into the preparation or execution of future science or calibration programs.

HLSPs are those derived at least in part from JWST (or other hosted mission) data holdings. These products generally include measurements of astrophysical target:

- locations
- brightnesses
- shapes
- timeseries
- spectral feature or light-curve characterizations

or other scientific attributes; or are:

- stacks,
- mosaics,
- or other combinations

of images or spectra that are not produced by the standard processing pipelines. HLSPs for JWST may be be composed or derived from non-JWST data, including ground-based data, provided that they are critical to the goals of the primary JWST program.

Requirements for HLSP contributions

In order to enable meaningful searches on contributed datasets, and to provide a consistent and comprehensive set of metadata to users, there are a variety of requirements and recommendations for contributed datasets, in the following categories:

- Acceptable data formats and data organization
- Metadata content
- File nomenclature
- Data provenance

See the specific guidelines for contributing HLSPs to MAST for details.

Contacts

MAST users who need assistance with archived data, or with the use of any of the related tools, should reach out to their Archive Contact Scientist (if applicable), or the MAST HelpDesk at archive@stsci.edu.

References

MAST Guidelines for HLSP Contributions
MAST YouTube videos of the Data Discovery Portal
JWST technical documents