



STScI | SPACE TELESCOPE
SCIENCE INSTITUTE

EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

JWST Pipeline Development

David R. Law
JWST Pipeline Lead

JSTUC Meeting
November 11, 2024



Update: JWST Pipeline (JP) Coordination Team

The JWST pipeline will never automate everything for everybody, but it should be capable of doing most things for most people.

In January 2024, STScI adopted a new framework for management and oversight of the JWST pipeline.

- JWST Pipeline (JP) Coordination Team has broad responsibility for all issues relating to the quality of the data products and documentation thereof being provided to the community.
- Set priorities, cross-division communication within STScI. Investigate, delegate, and track issues as necessary.
- Led by D. Law, with representatives from each instrument team.
- Common point of contact for questions and data-related issues.



Update: JWST Pipeline (JP) Coordination Team

The JWST pipeline will never automate everything for everybody, but it should be capable of doing most things for most people.

Updates since March 2024:

- Team has settled into a routine
- Weekly meetings (35 so far this year)
- Roles and responsibilities better defined, as is scope of concerns
- Deputy lead (Chris Hayes; NIRSpec) appointed October 2024





Overview of JWST Pipeline Builds

A 'build' is a released and tagged set of software including the core 'jwst' package and related dependencies.

See [JDox](#) for the latest table

Build lifecycle:

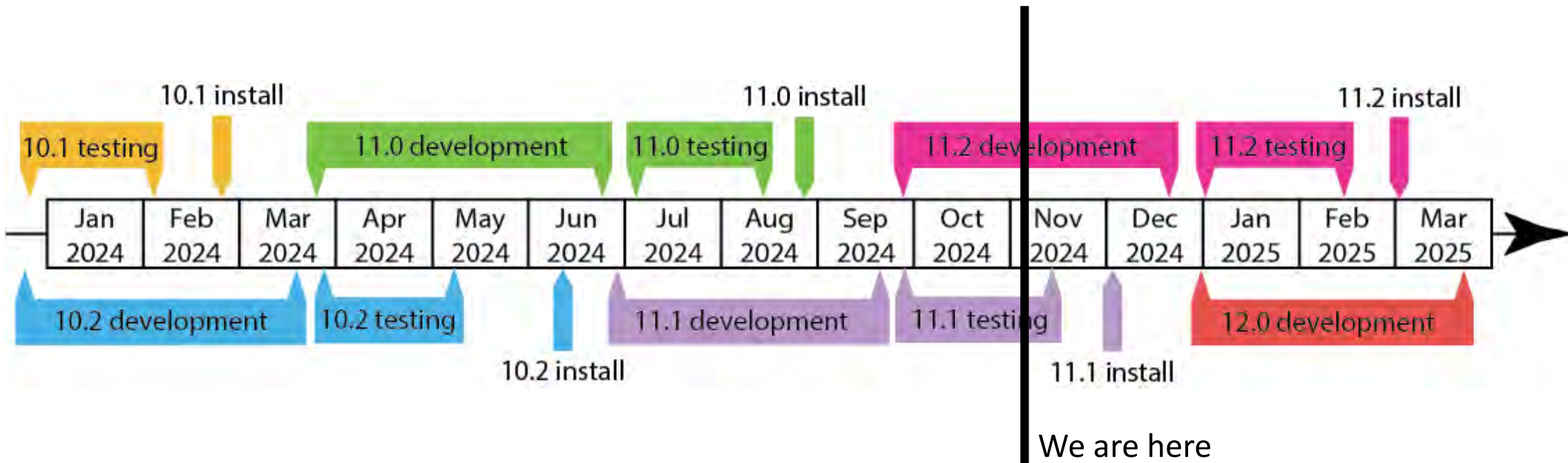
- **Development**
 - Changing daily, not recommended for non-developers. 'jwst' main branch.
- **Candidate**
 - Development finished, working on testing and documentation. Available for interested users.
- **Operational**
 - Default recommendation for users. Used to produce MAST data products.
- **Archived**
 - Old versions that are no longer current or recommended

Table 1. Release notes and associated versions for JWST Operations Pipeline releases

Operations Pipeline Build	Science Calibration Pipeline Version (jwst)	Build Status	SDP_VER	Candidate Released	Operations Installation	CRDS context
11.2	TBD	Development	TBD	Estimated Dec 2024	Estimated Mar 2025	Latest
11.1	1.16.0	Candidate	TBD	2024-09-20	Estimated Dec 2024	1295 (TBC)
11.0.1	1.15.1	Operational	2024.2.1	2024-06-26	2024-08-27	1293
10.2.1	1.14.1	Archived	2024.1.1	2024-03-29	2024-06-12	1240



Where are we now?





Major change in October 2024: Quarterly Calibration Updates

Reference files managed via 'CRDS' (<https://jwst-crds.stsci.edu/>) and versioned via context (e.g., `jwst_1290.pmap`).

Prior to October 2024:

- All versions of the pipeline always use the latest reference files by default
- This was a problem:
 - New files or pipeline parameters aren't always compatible with old software versions
 - *Example: New 1/f correction step couldn't be enabled because it didn't exist in prior versions*
 - Rerunning pipeline can produce different results from day to day or crash
 - Efforts to mitigate this led to lengthy delays in providing new software and ref files to community

After October 2024:

- Each build defaults to a fixed CRDS context; software and ref files tested together.
- Users should always get the same results when running a given build, and should be better able to reproduce MAST results.
- Reference file deliveries proceed as usual, but won't be default until next quarterly build.
- Allows far greater freedom in updating reference files and best pipeline parameters immediately



Major change in October 2024: Quarterly Calibration Updates

Reference files managed via 'CRDS' (<https://jwst-crds.stsci.edu/>) and versioned via context (e.g., jwst_1290.pmap).

- Change is retroactive to previous builds- applies to any **new** installations after October 7 2024
- Users should update pipeline software regularly in order to get the latest calibrations.
- Power-users can still use absolute-latest reference files by specifying a CRDS context, but files not guaranteed to be compatible with older pipeline versions
- Additional details are provided via [JDox](#) and [JWST Observer News](#)



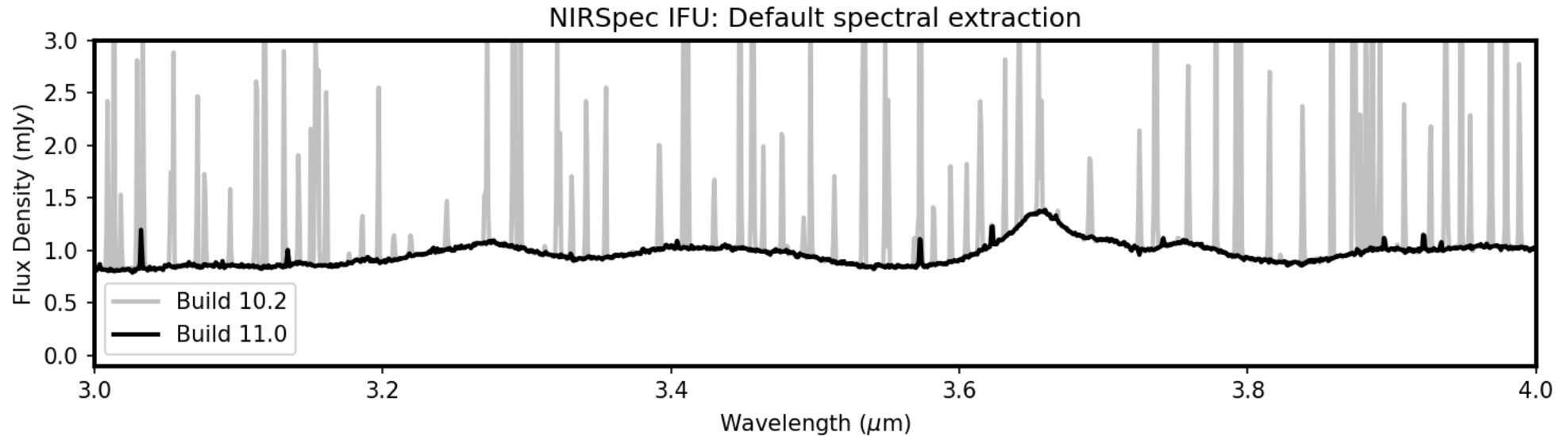
What has changed in the pipeline?

The JSTUC meeting in March 2024 highlighted six key areas of focus based on user surveys, data products workshop, and helpdesk tickets:

- 1) Bad pixels/cosmic rays/outlier detection
- 2) Correction for 1/f noise
- 3) Runtime (detector1 pipeline stage in particular)
- 4) Need better JDoX documentation of the pipeline data, known issues, and options for reprocessing for each instrument mode.
- 5) WCS alignment
- 6) ERR estimates



What have we done about outliers and bad pixels?

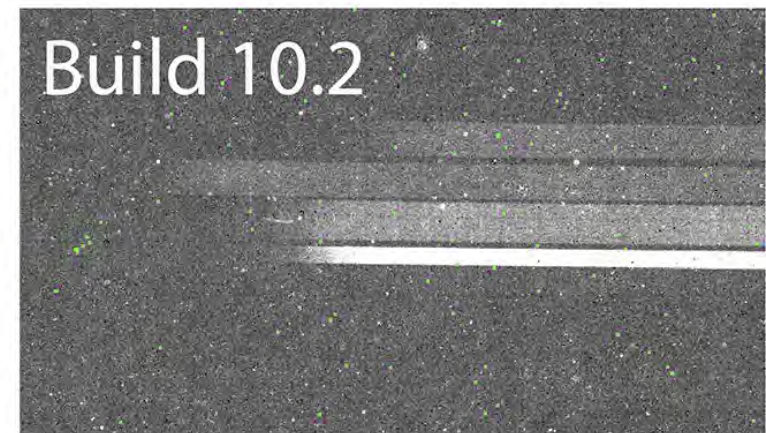
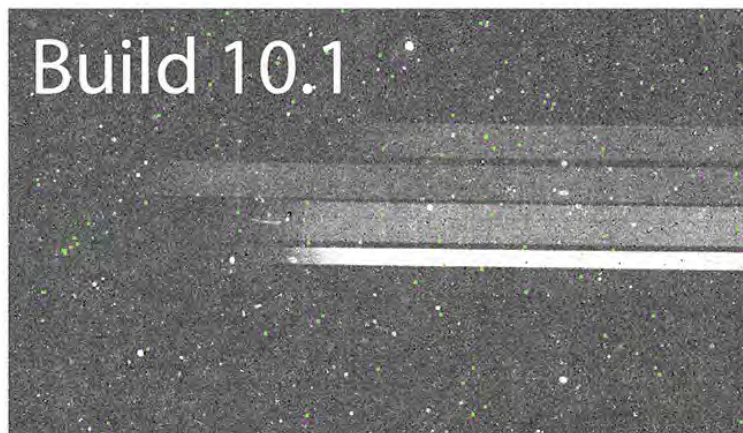
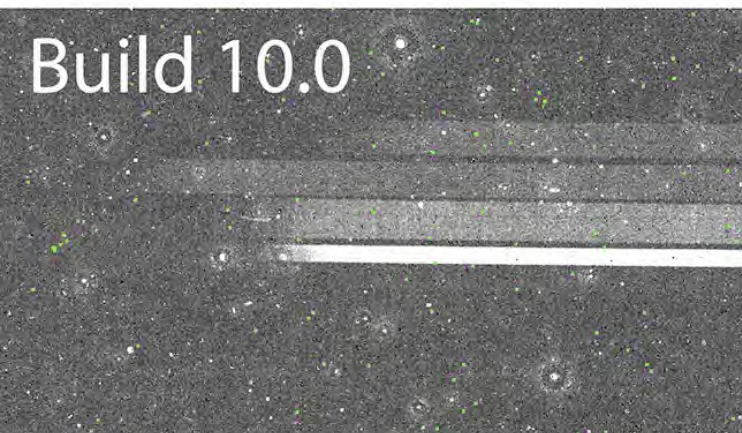
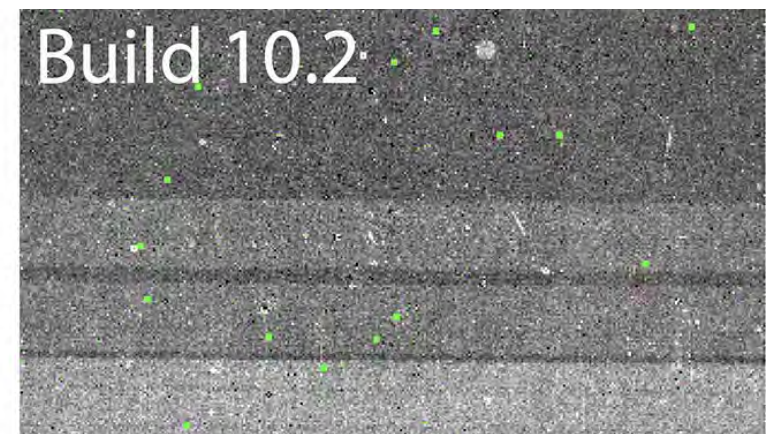
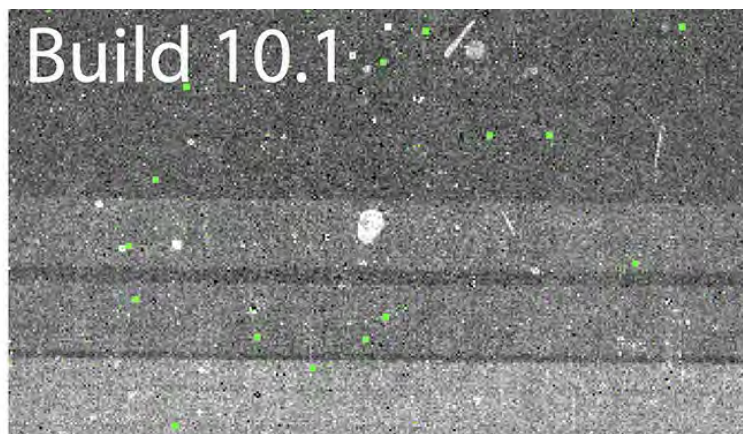
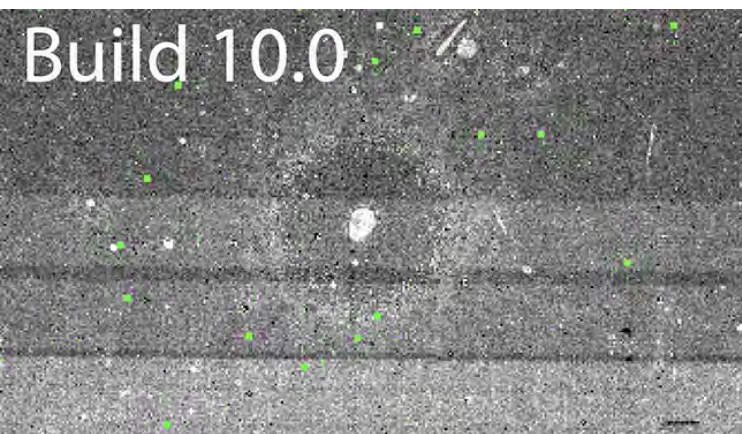


- Fixed unflagged cosmic rays saturating in the 2nd group for NIR instruments (see figure; B11.0/B11.1)
- Fixed NIRSpec IFU flatfield artifacts (B11.0)
- Improved detection of asymmetrical snowballs in NIR data (B11.0)
- Added new self-calibrated bad pixel masking step (B11.0)
- Moved pixel replacement to spec3 to fill in detected outliers (B11.0)
- Improved consistency between data quality flags and SCI values (B11.1)



What have we done about outliers and bad pixels?

(Mike Regan and Eddie Bergeron)



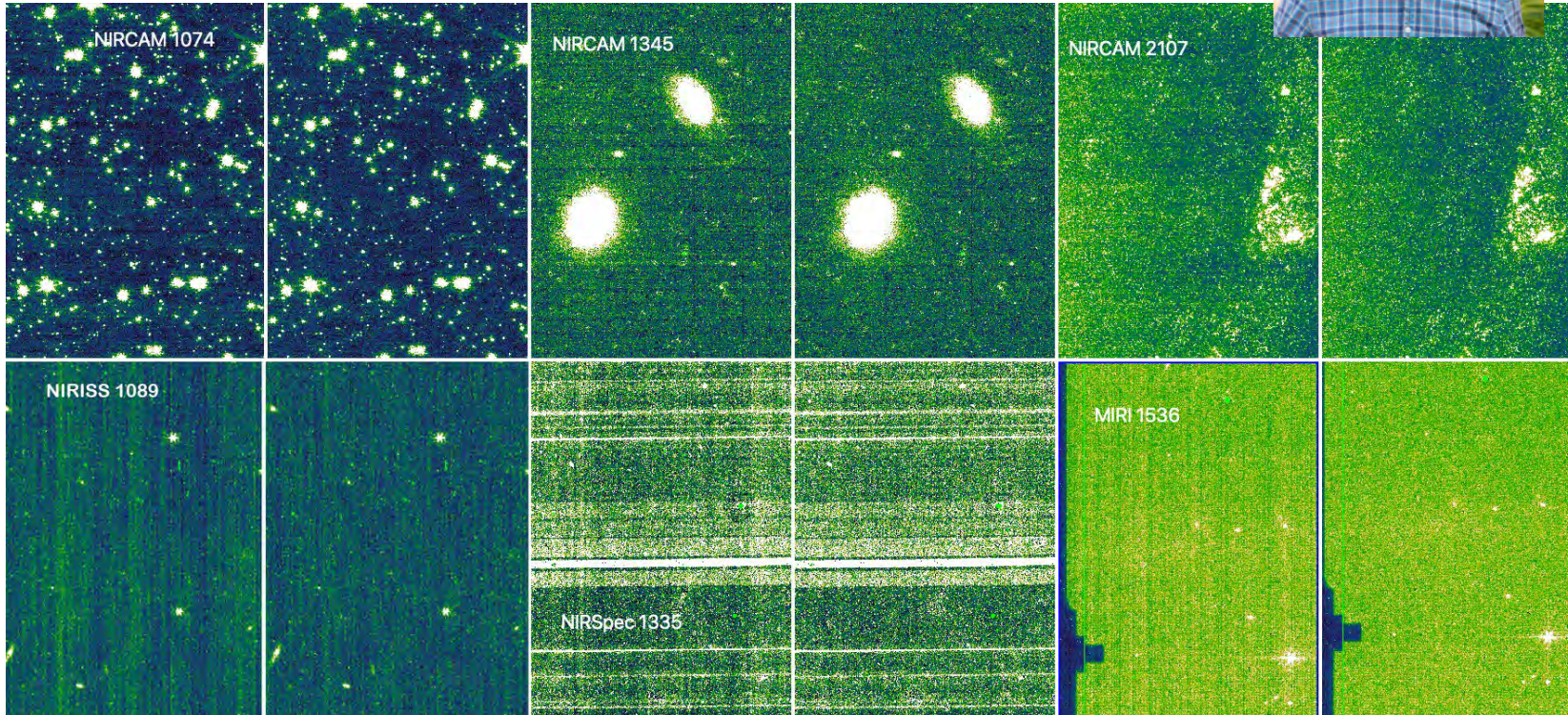
- Improved detection of snowball persistence (B10.2) and asymmetrical snowballs (B11.0) in NIR data



What still needs to be done about outliers and bad pixels?

- Improvements for NIRSpec MOS and fixed-slit data (B11.2)
- Improvements to MIRI cosmic ray shower detection, particularly for MRS, and enable by default (B11.2)
- Further testing of self-calibrated bad pixel step to make available through pipeline notebooks (B11.2)
- Enable default pixel replacement for additional modes beyond MIRI LRS (B11.2)
- Time-dependent bad pixel reference files for all instruments, confirm that reference files identify them properly
- Improve handling of telegraph pixels in NIR instruments

What have we done about $1/f$ noise?



- Implemented a general-purpose routine in Detector1 to correct $1/f$ noise for all JWST NIR instruments ([clean flicker noise](#))
- Two community-based algorithms: linear (Chris Willott) and fourier (Bernie Rauscher; NSClean). Revised old NSClean to leverage the new code.
- Major work by Tim Brandt and Melanie Clarke



What still needs to be done about **1/f noise**?

- Step is available, but currently off by default pending further testing (wide range from deep field imaging to nearby galaxies and NIRISS WFSS)
- Aim to begin enabling 1/f correction for some modes in B11.2 and beyond if testing indicates can be automated
- Implement routine to fix NIRSpec thermal picture frame artifacts (prototype in progress)



What have we done about pipeline runtime?



- **Rewrote the time-consuming ramp fitting step in C**
 - Achieves factors of up to 200 improvement in runtime (B11.0, B11.1)
 - Some long MIRI exposures reduce from > 1 hour to < 20 seconds
- **Improve runtime of the NIRSpec IFU pipeline (B11.1)**
 - Spec2+spec3 runtime reduced by 60%
- **Improve runtime of the NIRSpec BOTS/MOS pipeline (B11.1)**
 - BOTS spectral extraction speed improved by factor of 5x
- **Improve runtime of the NIRSpec NSClean step for BOTS (B11.1)**
 - Factor of 30-100x speed improvement
- **Major work by Ken Macdonald and Tim Brandt**



What is still to be done about pipeline runtime?

- Improve runtime of the jump step in calwebb_detector1, which is now the tall pole for total runtime.
- Improvements to MIRI EMICorr runtime
- Further improvements to NIRSpec runtimes



What have we done about **pipeline documentation?**

- Restructured JDoc pages presenting all information about the [pipeline](#), [Known Issues](#), and [Calibration Status](#) (March 2024)
- New [jwst-pipeline-notebooks](#) repository for end-to-end processing
 - All notebooks follow a similar layout, process from uncal files to Stage 3 products
 - Minimalist, but provide example of how to modify various steps
 - Existing notebooks for MIRI MRS (template) and NIRISS Imaging
 - Final checks of MIRI imaging and NIRCcam imaging
 - Eventually notebooks for all instrument modes
- **AAS JWST pipeline workshop scheduled for January 2025**
 - Presentations on pipeline structure and overview, known issues and workarounds
 - Walkthrough of installation and usage of new jwst pipeline notebooks on demonstration flight data
 - Hands-on support for user pipeline issues (demo and regular science data)
 - ~20 people registered so far



What still needs to be done about pipeline documentation?

- Cleanup of the ReadTheDocs documentation
- Better guidance for users on which optional parameters to tweak and why on the instrument mode pages.
- Considering another JWST pipeline workshop once all of the changes since the workshop last Fall have percolated throughout science analyses



What have we done about **WCS alignment**?

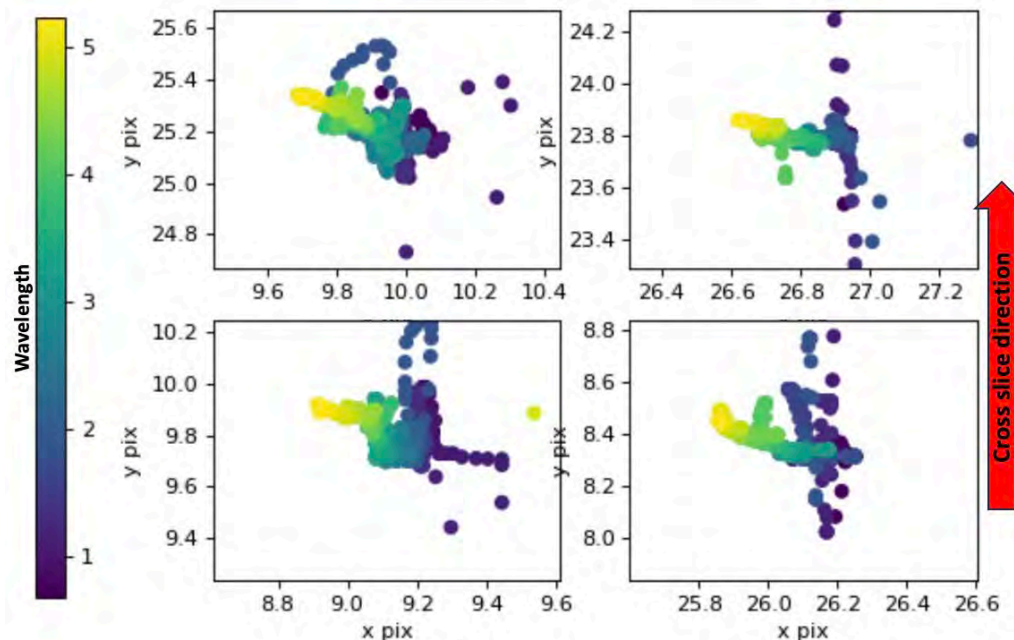


- Updated SIP polynomial approximations for imaging data to be more representative of the full WCS solution
- Updated FGS2 geometric calibration to fix ~ 0.1 - 0.2 arcsec offset with respect to FGS1 (spearheaded by Tony Sohn; some pipeline fixes still required)
- Updated guide star catalog coordinate reference data
- Updated algorithm and parameter reference files used by the tweakreg step for imaging (B10.2/B11.0)
- Update cube building software to allow users to refine WCS of individual input exposures (B11.1)

What still needs to be done about WCS alignment?



- Investigation underway into issues with NIRSpec IFU WCS solution
 - Inaccuracies in spatial and spectral distortion solution
 - Source 'wander' with wavelength
 - 'Rotation curves' for point sources
 - Spearheaded by Tracy Beck



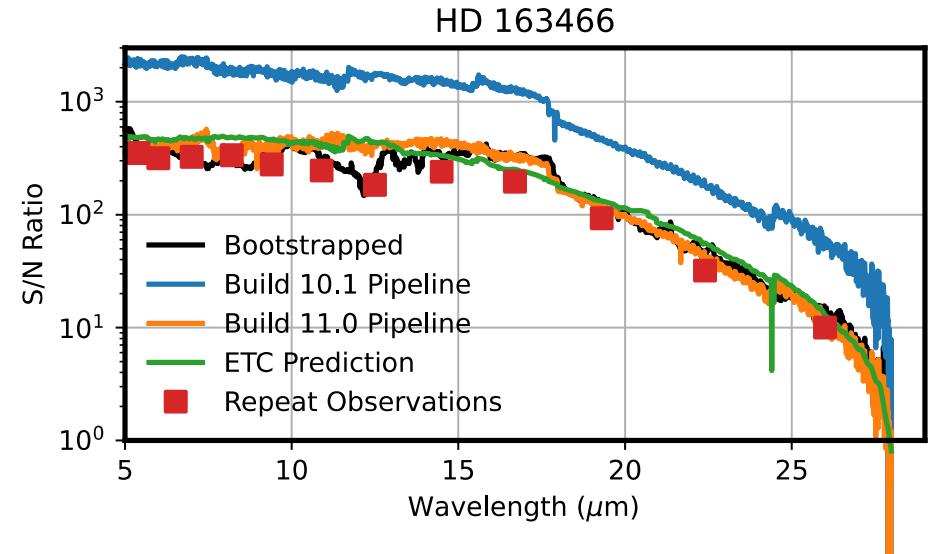
NIRSpec prism source position vs wavelength (Jdox; T. Beck)

- Add automatic WCS adjustment to MIRI MRS point source data to account for grating wheel repeatability



What have we done about **estimated uncertainties**?

- Added term for Poisson noise on the dark that can limit deep observations (B11.0)
- Multiple bugfixes to error propagation throughout the pipeline, including better handling of NaN values (B11.0/B11.1)
- Added treatment of IFU covariance during MIRI MRS spectral extraction (B11.0)
- Updates to MIRI MRS flatfield uncertainties





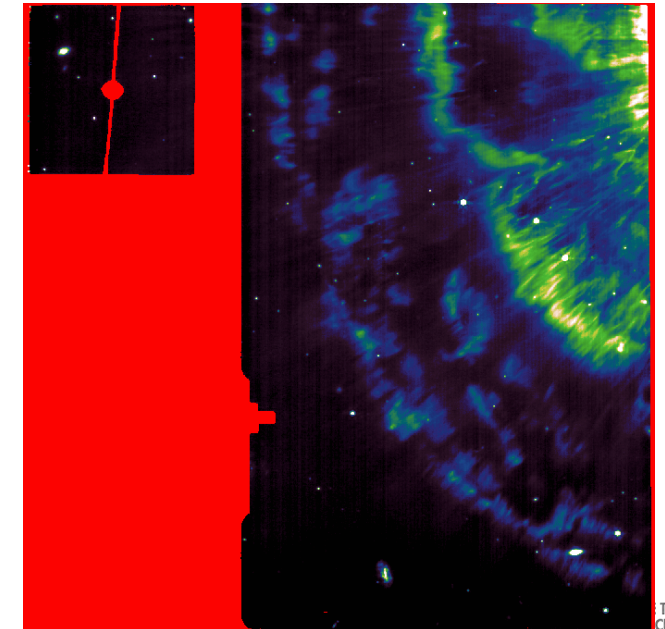
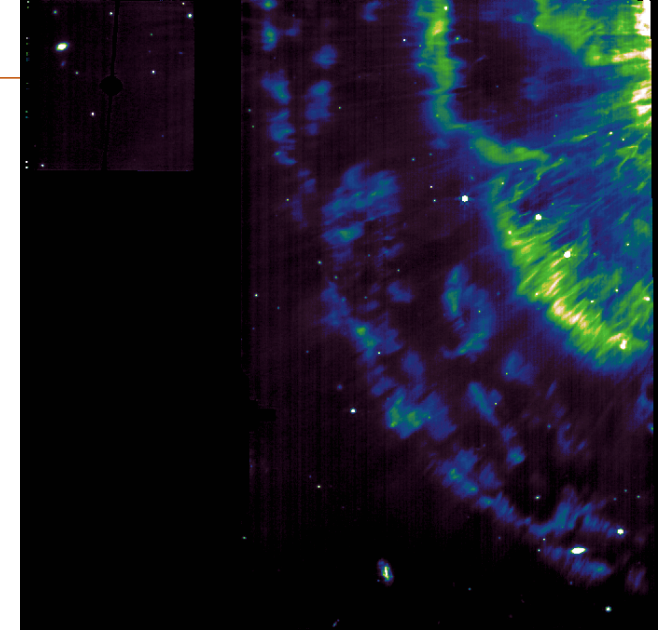
What remains to be done about **estimated uncertainties**?

- Refactor drizzle code to be able to properly resample variance arrays (limitation dating back to HST- currently workarounds have limitations)
- Enable Poisson dark errors for NIR instruments
- Systemic analysis of performance vs ETC for all instruments



Other major developments

- Implemented a system for instrument team regression testing of pipeline data products for each build.
 - Manual for Builds 11.0, 11.1, working on automation
 - Ensures instrument team eyes on representative end-science data products
- Changed handling of out-of-field information
 - Empty areas of the FOV are now NaN, not 0 valued
 - Makes clear where there is no data
 - Ensures that regions without data are not being processed as if they had valid data





Other major developments

- **Refactoring data handling methods throughout pipeline**
 - Primary aim is to improve memory management
 - Integrating testing of runtime and memory pressure to development workflow
- **New likelihood-based ramp fitting algorithm (Tim Brandt) aimed at optimizing performance in deep fields and extremely short ramps (In Testing)**
- **Overhauled spectral resampling for NIRSpec MOS and FS modes**
 - Significantly better performance
 - Led by Melanie Clarke and Chris Hayes
- **Improved NIRISS SOSS spectral extraction**
 - Folds standalone PASTASOSS package into the pipeline
 - Led by Joe Filippazzo and Tyler Pauly



Optimizing NIRSPEC deep field observations

Working with science teams to fold lessons learned back into the pipeline

- E.g., GO-2123 (PI Kassin): GARDEN team obtained ultra-deep stepped MSA spectroscopy of high-z galaxies probing faint-source regime
- Eddie Bergeron, Mike Regan, Jane Morrison particularly critical to this effort

Some key areas of potential improvement (incomplete list):

- Picture frame correction (due to thermal instability of detectors)
- Even/odd row correction (due to transient bad pixels in IRS2 ref pix)
- $1/f$ noise (already in pipeline; using as a test case)
- Improved dark reference files
- Improved barshadow corrections
- Sporadically fail-to-open or fail-to-close shutters contaminate background est.



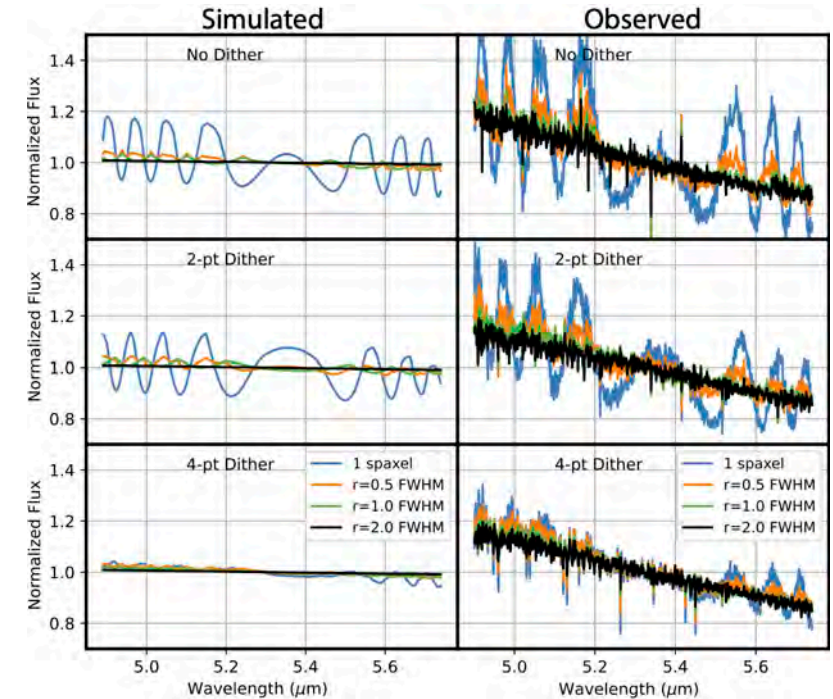
What else are we working on now?

- Adding log to output data model
- PSF-based spectral extraction (and general overhaul of 1d extraction)
- Improvements to background subtraction for NIRSpec MOS, NIRISS WFSS, MIRI MRS
- High-level review of major science modes and optimizations to benefit the community
- MRS fringing
 - Provide per-band cubes and both regular and residual-fringe-corrected 1d spectra by default.



What other common user issues are there?

- Oft-reported issue: resampling artifacts in IFU data cubes (NIRSpec+MIRI)
 - Natural consequence of building rectified cubes that resample data that was not Nyquist sampled (see Law+23)
 - Nonetheless a big annoyance for users used to doing spaxel analyses
 - Community is trying to develop alternatives, which can work ok for some science cases
 - Should we put resources into this?



IFU undersampling (Law+23)



Summary

- Making progress in all major areas
- New framework for pipeline management is working well, and responsive to user feedback (some reported issues fixed within hours)
- High-priority items continue to be worked in the upcoming months
- Long-term optimization plan under development
- Upcoming pipeline workshop at January AAS meeting
- Will continue to incorporate feedback from user community into prioritization.