Science Data Products

Learning which of the many JWST data products are appropriate for science analysis requires a basic understanding of how the data are processed, the file naming scheme, and the format and semantic content of the science-ready products.

On this page...

- Science Data Processing
  - Level vs Stage
- File Naming Scheme
  - Exposure- vs. Target/Source-based File Names
  - File Contents
- For Further Reading...

Science Data Processing

A key point to understand is that, with *every* JWST Observation (e.g., a row in the Portal results table), *all* science and associated data products are bundled with it. These products include:

- All calibration levels of science data processing (see the summary in the Level vs Stage subsection below)
- Various supplemental products

In all, the set of files per Observation could number in the hundreds or even thousands. Exactly which calibrated products are generated depends upon the configuration of the instrument that obtained the data (imaging, various forms of spectroscopy, coronagraphy), and the operating mode (time-series, dithered, moving target, etc.). See the detailed description of Science Products for more information.

The other products are useful either for evaluating the quality of the data processing, or to provide inputs in case you wish to re-process data with some or all stages of the calibration software on your own computer platform.

When searching for products in the MAST Portal, those identified in the MAST Portal Download Manager as Minimum Recommended Products (MRP) are a good approximation to the most relevant for science analysis. These files usually number less than a dozen, except for multi-object spectroscopy where there is one highest-level, combined, extracted product per spectrum (i.e., potentially hundreds of spectra).

Level vs Stage

You will encounter slightly different labels to denote Science products, depending upon the document collection you are reading.

- **JDox:** The JWST mission document collection focuses on the pipeline processing ***stages*** that are used to produce different products. See Stages of Processing for details.
- **Processing (DMS):** The ***data management system*** (which processes all JWST science data) focuses on the ***Level*** of data product that is being produced; for products visible to users this starts with L-1b and ends with L-3. For comparison, the Stage 0 pipeline takes DMS L-1b products and produces L-2a products, which themselves comprise the input for the Stage-1 pipeline, and so on.
- **Archive (MAST):** The MAST Common Archive Observation Model (CAOM) database labels products by a generic calibration level that applies to many instruments and telescopes, whether ground- or space-based. It does not currently distinguish between DMS L-2a (slope corrected) and L-2b (instrument signature removed, calibrated to physical units).

The following table summarizes the products and the translation between the vocabularies. You may also find helpful the diagrams in the Linkages in the Portal article.

Description	JDox Data Stage	DMS Processing Level	MAST/CAOM Calibration Level
Planned/approved but not yet executed observations	N/A	N/A	–1
Uncalibrated FITS files with all header keywords populated, per detector/exposure	Stage 0	L–1b	1
Slope-corrected FITS files (up-the-ramp slope removed), per detector	Stage 1	L–2a	2
FITS files, per detector and exposure, where instrument signature has been removed and the data have been astrometrically and radiometrically calibrated	Stage 2	L–2b (also L-2bs, L-2c, L-2cs)	2
Combined, calibrated science product per target or source	Stage 3	L–3	3
Contributed Science Products	N/A	N/A*	4

*DMS does produce L-4 products for HST, but presently no such products are defined or produced for JWST.

The target/source-based combined products (L-3), and the fully calibrated exposure-based images from which they were derived (L-2b), are the most likely to be relevant for your science analysis.

File Naming Scheme

The file name convention for JWST data products was designed in large part to facilitate automated processing of data products per observation. The names are not scientist-friendly, and can be as long as 70 characters. It is useful to learn enough about the naming scheme to identify products that are relevant for science analysis. In particular, L-3 product filenames are distinguishable from all the others, as explained below.

Exposure- vs. Target/Source-based File Names

Lower-level, *exposure*-based products correspond to a particular program, observation, visit, instrument, detector, and exposure. The file names are similar in structure to:

```
jw00689001001_02201_00001_nrcalong_uncal.fits
```

Note the program ID (`00689`), observation number (`001`) and visit number (`001`) concatenated near the beginning of the name. The string following the last underscore is the data product semantic type.

Higher-level, *target*/*source*-based products have file names similar in structure to:

```
jw00689-o001_t001_nircam_clear-f356w_i2d.fits
```

Note the program ID, followed by a dash, and then additional information including the instrument name and optical elements. See the general discussion of the File Naming Conventions article for background, and the definitive File Naming Schemes page for details.

File Contents

The semantic content of the files is encoded in the filename suffix (the part of the filename just preceding the file extension). For example, filenames ending with `_uncal.fits` are uncalibrated (Level-1) FITS-formatted data, while a file ending in `_cat.ecsv` is a Level-3 source catalog in ECSV format. See MAST Data Product Types for a brief description of all semantic types found in MAST. See also JWST Data Product Types or Pipeline Suffix Definitions for the full set of suffix names, and Science Products for a detailed explanation of pipeline product contents and internal organization.

Most, but not all, data products are in FITS format (FITS 4.0 Standard). See Data Product File Formats for a summary of data products and their formats, and the article on Data Product Types to learn about the detailed organization and content of each product.

For Further Reading...

The on-line documentation in JDox and in the calibration documentation (CAL) is very detailed and comprehensive.

- JWST Data File Naming Conventions (JDox)
- JWST File Naming Schemes (CAL)
- Data Product Types (CAL)
- Data Product File Formats
- Science Products (CAL)