What Calibration and Imaging products will be delivered to me?

Catarina Ubach - 2023-04-13 - Archive & Data Retrieval

What you see in your data package when you download it from the archive will depend on whether the imaging and calibration have been performed by the pipeline, or one or both via a manual process.

In this article we give a quick guide to the data packages, and more detailed information on the imaging pipeline products for data processed after October 1 2016.

**Full description of delivered products can be found at:**


**Other relevant references:**

- **What is Remcloud, and how could it reduce phase RMS?**
  [https://help.almascience.org/kb/articles/what-is-remcloud-and-how-could-it-reduce-phase-rms](https://help.almascience.org/kb/articles/what-is-remcloud-and-how-could-it-reduce-phase-rms)

In all cases, there are five sub-directories at the lowest level of the directory tree (corresponding to the member observation unit set, MOUS), plus a README file:

- **calibration**
- **log**
- **product**
- **qa**
- **script**
- **README**

The **calibration** directory contains calibration and flagging tables that can be applied to the raw data by **scriptForPI.py** in the **script** directory. In the case of pipeline imaging, the **calibration** directory also contains the file used to define the continuum for continuum subtraction, `<MOUS UID>.cont.dat`.

The **log** directory contains the logs of the CASA runs (manual or pipeline) used to calibrate the data and make the image products. For pipeline runs it also contains a file `<MOUS UID>.cont.dat`. 
The `casa_commands.log` file contains a list of CASA commands that can be used to rerun (parts of) the pipeline process.

The `qa` directory contains QA information. For pipeline calibrated data, a weblog is included that can be unzipped and unzipped and read by a browser. If the imaging pipeline has been run, the weblog includes the imaging steps. For manual calibrations, the QA information is in several files that together constitute a QA summary for the data.

The `script` directory contains a set of scripts that can be used to recreate the delivered projects. Details on how to run the scripts to reproduce the pipeline products are given in the User’s Guide in the ALMA Pipeline Documentation available at [http://almascience.org/documents-and-tools/pipeline-documentation-archive](http://almascience.org/documents-and-tools/pipeline-documentation-archive). Note that, in the case of pipeline imaging, the `scriptForImaging.py` is a null script. For manually calibrated data, scripts for each execution can be found in the files `<Execution Block UID>.scriptForCalibration.py`.

The `product` directory contains the image products. The number of products supplied depends on the process used. Fully manual calibration and imaging makes a subset of products tuned to the PI request, usually at the “bandwidth for sensitivity” picked by the PI in the OT. Similarly, datasets that have been imaged manually contain these tuned products, and also images of the phase and bandwidth calibrators made by the pipeline. Data that have been imaged by the pipeline contain a standard set of image products, including continuum and a cube for each spectral window, made at the native resolution in the LSRK velocity frame. Details of the image products are given below:

**Pipeline imaging products**

**Pipeline products**

By default, the pipeline will produce a full set of products (aggregated continuum, plus continuum and cube images for every science spectral window) for every source. However, in practice only a subset of data may be imaged if the computational resources required to make a full set of products are too large. The images are released once a QA2 assessment (Pass or SemiPass) has been made. As the pipeline heuristics are still being improved, in some cases you will notice problems with, for example, the continuum subtraction, insufficient clean depth and/or masking, and/or images that lack self calibration (which is currently not implemented in the pipeline). The CASAGuide [ALMA Imaging Pipeline Reprocessing](http://almascience.org/documents-and-tools/pipeline-documentation-archive) outlines how you can improve upon the imaging from the pipeline.

For a comprehensive guide to the imaging pipeline, please refer to the [ALMA Pipeline User’s Guide](http://almascience.org/documents-and-tools/pipeline-documentation-archive). Bear in mind that most ARCs and ARC nodes encourage data reduction visits if you would like help with improving your images.

**Naming convention for pipeline products**

Image products are named according to the following convention:

`<MOUS UID>.<Source Name>_<intent>.<spectral`
MOUS UID is a string that uniquely identifies the dataset (e.g. uid___A001_X88f_X270)

<intent> is the observation intent of the source, e.g. sci for science target, ph for phase calibrator, bp for bandpass calibrator.

The spectral window list contains the spectral window numbers used in the product (e.g. spw17, spw17_19_21_23). The quickest way to identify which spectral window corresponds to which frequency/resolution combination is to click on the link to the measurement set on the Home page of the weblog, then to click on the "LISTOBS OUTPUT" button.

<dimenstonality> is either mfs for multifrequency synthesis (resulting in an image with two spatial dimensions), cont for continuum aggregated over all spectral windows (two spatial dimensions), or cube for a cube with two spatial axes and a frequency/velocity axis. You may also see tt0 and tt1 for mfs images made using the the zeroth and first Taylor terms, respectively. The tt0 image corresponds to the regular image, the tt1 image is related to the spectral index image.

<imgetype> is pbcor for a primary beam corrected image, or pb for the primary beam image.