**Information from the on-line review site:**

<https://pdssbn.astro.umd.edu/howto/review_feedback.shtml>

Minimum Required Feedback

Documentation

* Does the dataset contain all documentation needed to use and understand its data without prior knowledge?
* Is the provided documentation well organized, clear and self-consistent?
* Can the dataset be understood without any external documentation it references, or should the information in said external references be incorporated into the dataset?
* If reviewing calibrated data, does the documentation fully explain the calibration process and contain all necessary parameters needed to repeat it?

PDS Labels and Meta Data

* Are the descriptions and scientific content contained inside the PDS labels sufficient to understand their corresponding data products?
* Is all significant meta data included directly in the PDS labels?
* Do the labels provide all essential description of data values directly in the label, instead of deferring them to external references or documentation?
* Can the data be read programmatically using only the information contained in the PDS labels?

Data

* Does the data look physically reasonable when examining it by eye or via a display tool?
* When displaying the data as plots or images, are there any unexpected deviations?
* Formulate a scientific inquiry and attempt to use the data to answer the inquiry.
* If reviewing both raw and calibrated data, attempt to calibrate a raw data file.

Reviewers should remember that they are reviewing a PDS archive submission. All documentation, calibration observations and other ancillary information needed to understand and use the data should be present in the archive. Additionally, an important part of the review is ensuring that the label describes the data both functionally and scientifically. Please use either PDS Tools (*see box on right*) or other software written specifically to read the data according to the PDS labels for at least part of your review.

**Example issues found by reviewers**

* Unexplained gaps in the data (missing files in a series, for example).
* No pointing information included in the label for a space mission.
* A filter number is given, but the filter's center wavelength and bandpass are not provided in the label.
* The results obtained from the data during the review are different from results presented in a publication that used the same data.
* While looking for necessary information about the data, the value was not contained in the PDS labels, instead being found in a FITS header.
* A table column named Observation Time does not describe the time system it is using, nor whether it's recording the start, midpoint or some other point in the observation time.
* It is unclear which calibration parameters are applied to what data.
* The observed object could not be found in the image.
* It is documented that external conditions may have affected some results, but not clearly indicated which results may have been affected.
* The first few rows of a table look physically reasonable, but the rest appears to be noise.
* The data are for a comet flyby, but the actual time of the flyby is not found in a high level overview of the dataset document.
* Units of measurement are not included in the labels.
* The provided log of observations is inconsistent with start and stop times in the labels.
* A figure referenced in a document or label could not be found.
* High resolution observations taken adjacent to low resolution observations show dramatically different results, without any explanation in the documentation.

**Instructions for reviewing ROSINA data**

From PSA to the ROSINA reviewer:

Let me add some notes on what you will find in the datasets and what we want you to focus on.

You will see that there are three types or levels of datasets:

  \*  RO-C-ROSINA-3-XXX-V1.0 = Level 3= Calibrated data (according to PDS definition of L3)

  \*  RO-C-ROSINA-4-XXX-V1.0 = Level 4= Resampled data (according to PDS definition of L4)

  \*  RO-C-ROSINA-5-XXX-V1.0 = Level 5 = Derived (according to PDS definition of L5)

Within each level there are many datasets covering different mission periods. As there is so much data then you might want to choose data for each level from different

time periods in the mission in order to cover more periods without multiplying the work.

A summary of what you will find in each dataset and what we would like you to focus on is given below:

Level 3

- Should include data for the whole comet phase including RTOF and DFMS (COPS is included in Level 4 only)

- Should include fragmentation and sensitivity tables.

- The MTP8 directory of the RO-C-ROSINA-3-PRL-V1.0 dataset contains H2O DFMS spectra with commanded mass 18 with better precision in the mass scale (L3 improved product –

sample of ~1 month for mass 18). Please, look at this particular dataset.

- A preliminary sample of Level 3 data including only RTOF data had already been reviewed whereas DFMS data is totally new. The calibration, specially the DFMS one, has to

be reviewed.

Level 4

- Includes calibrated COPS data for the whole comet phase.

-These datasets are new and have never been reviewed.

Level 5

- At this point in time Level 5 should contain times series (local densities) for all major species (H2O, CO, CO2, O2). In the future these datasets will contain other

minor species.

- These datasets are new and have never been reviewed.

Outside the dataset folders

In addition look to the Science User Guide draft document that can be found outside of the dataset folders. A final version of this document will be integrated into the

datasets in the future.

**Comments**

1. General
   1. As the ROSINA team attests, it takes a while to get comfortable with the data, but the extensive documentation makes that comfort possible.
   2. The user guide is enabling; thanks.
   3. Good: appear to have all the data and the calibration information at the PDS. Also have a lot of documentation
      1. The amount of work is mind boggling.
      2. Not so good: datasets are unwieldy, with the data in millions of different files with no information in their filenames except instrument and time.
   4. In the time that I had available, I could not hope to replicate the skill that the team developed in calibrating the raw data and developing the products, but I could review the documentation, its organization, and spot check some of the data. I found no non-trivial problems. In this context, “trivial” means that the problem is not an error and that the problem is only an inconvenience to users of the data.
   5. Suggest that the data user manual provide the initial sequence or steps to understand the data (see below).
2. Aareadme.txt
   1. The aareadme.txt file states that there are additional NMS documentation at the ESA servers and at the PDS node; these documents should be included in the references.
   2. The location of referenced documents such as OPERATION\_LOGBOOK.PDF should be identified.
   3. Is Kathrin Altwegg the correct contact person? I had heard that she retired or soon will.
3. Data user manual
   1. Good for the difficult part: description of the calibration and cautions. Could use more documentation on basic use of the database: where to find documentation, how the data files are organized, naming convention for files, etc.
   2. Where are naming conventions defined, e.g., ce\_20140425\_092001\_3\_m0160.tab
   3. There are many useful documents referenced in the manual; it would be useful to have their locations identified.
      1. For example, Sensitivity and fragmentation calibration of the ROSINA Double Focusing Mass Spectrometer has excellent information on the instrument, but the file name, th1\_dfms\_mh.pdf is obscure.
      2. Example: “the ROSINA User Manual including appendix D1-DFMS instrument modes” is said to be in “Part of the archive,” but the location is not defined (it is the document directory) and it identity is not clear from the filename (ad1\_inst\_op\_dfms.pdf), particularly when most of the many documents in the document directory have similar filenames.
      3. Users could be directed to docinfo.txt in the document directory, for a description of each document in the directory.
   4. Would be useful to describe approach to the data
      1. Begin with this guide
         1. Data organization
         2. Instrument description
         3. Data cautions
      2. Go to any of the datasets
         1. Get the directory structure from the aareadme.txt file.
         2. Get the operations\_logbook from the documents directory and use it to find the desired data period, which directs users to the correct dataset: first the major group (PRL, ESCn, EXTn) and then individual datasets (mtpnn).
      3. Use dataset.cat files in each dataset to locate the files that define the format for each data file.
   5. The operational modes are listed in \*\_mode\_desc.asc; where are they described?
   6. Could use another round of editing
      1. Some missing figures
      2. Some additional explanation of tables and figures.
      3. Some acronyms not defined.
      4. Location of some reference material could be more specific than “Located with the archive.”
      5. Some terms used in the dataset are not defined. For example, the CE or CEM dataset part of DFMS.
4. The dataset.cat files are useful.
   1. For 3-PRL files, the dataset.cat file could reference the aareadme.txt file for a description of the directory structure in the /data/ directory.
5. The aareadme.txt file describes a CEM detector for the DFMS and the CE data. Those acronyms are not defined in the file. Nor are those acronyms used or defined in the data user manual.
6. In EAICD, the equation for t0 in Section 2.5.3.4 is incorrect.
7. PRL 3 DFMS mtb2 CE
   1. All the files say quality 4, which means insufficient data to calibrate; is that correct? Some seem to have good signal.
   2. Mass calibration for the GCU data (mode 160) seems to be low by 0.05 to 0.1 u; is that expected?
   3. Mtp8 was supposed to have higher precision. Did not see this (still only two significant figures)
8. In the DFMS MC files, I do not understand the list of molecules.
9. Only reviewed one of the mc files as they took too long (hours) to download (only downloaded the smallest, 1.9 GB, and that took 40 minutes).
   1. This one, small dataset had 61,000 DFMS MCP files. (only reviewed 20,000 of them).
   2. Some have data quality 0 (best), and others 4 (these have only a few counts)
10. For the high-resolution data that are only in hundredths, are the values rounded or truncated? (Truncated works well.)
11. Some information, such as the descriptions of modes, is in more than one place (dfms\_mode\_id\_table\_\*.tab in calib directory and dfms\_mode\_desc.asc.txt in document directory), which is great.
12. rtof\_mode\_id\_table\_\*.tab (in ESC4 calib directory) references FRAG00 and SENS00, but there is only rtof\_frag01\_\*.asc in the directory: is “01” instead of “00” and does not say “sens”.
13. For column data, particularly calibration, mode, and other reference data, would be great to have one or more header rows so that users do not have to use the fmt file as often.
14. RTOF data (ESC4 SS) have many counts that are -1; is this intentional? If it is, it should be in errata or readme. Suggest that 0.0 is better indicator of problem than -1.
    1. OS do have zero as minimum.
    2. Me: read errata and readme docs
15. Are any of the RTOF data files quality 0 or 1? My spot check only found quality 4 (gas-calibration files are quality 0). How does a user find the higher-quality data?
16. DFMS H2O looks good
    1. smooth variations
    2. reasonable comparison to COPS (it is slightly higher than H2O at all measurements)
    3. Much higher than CO2 density (another file)
17. In DFMS level 5, could add uncertainty for the density provided by COPS.
    1. In detailed examination of one file, saw x2 difference in COPS density from H2O and CO2 file, but that could be within the error 2e13/m3. (But those two densities usually within 5% of each other.)
18. RTOF (EXT3, mtp32): water uncertainty is fixed 20%.
    1. Is rationale described somewhere (check aareadme, etc.)
19. RTOF (EXT3, mtp32):
    1. Water is only 0.2 to 0.5 of the COPS density, with lots of CO and CO2, too.

**9 October 2018, Notes from review, US review.**

* I did not use PDSview.
* Did not review all of the documentation, only what I needed to understand the data organization and primary characteristics.