Policy for Use of Compression in PDS Archives

Adopted 2005-10-17
Amended 2005-11-14
Amended 2013-11-19

This policy applies only to data archived in PDS3.

Data may not be archived in compressed form except as follows.

1. Large volumes of data received from spacecraft in compressed form may be archived in that compressed form subject to the following conditions:

    a. The decompression algorithm must be non-proprietary.

    b. A detailed decompression algorithm (or reference to a detailed algorithm in published literature) must be provided as part of the archive.

    c. Software source code for decompression must be provided in at least one high-level programming language in common use by the science community (e.g., C, C++, Java, Fortran, or IDL). The source code is intended not as operational software but as a "skeleton" program that can be adapted to new computer systems and operating environments. The source code captures the subtle implementation considerations of the compression algorithm that may not be apparent from the algorithm description.

    d. Before and after examples of the data decompression algorithm must exist in the archive for testing software implementations of the decompression algorithm.

2. With explicit permission from the appropriate Discipline Node, derived image products may be archived in compressed form under the following conditions:

    a. The source version of the image products must be archived in uncompressed form.

    b. Compression must be lossless unless the lossless requirement is explicitly waived by the Discipline Node.

    c. The compression format must be approved in advance by the Discipline Node and the Management Council.

    d. Product meta-data must identify the compression algorithm (or software) used in compressing the product, and its version

    e. All PDS meta-data for each product must be available in uncompressed form

    f. The PDS must have a copy of the specification or standard defining the compression algorithm used, at the version level that was used. If legally permitted, the documentation should be included in the archive.

g. Decompression software must be capable of producing a correctly formatted and labeled decompressed PDS data file. Additional output formats are permitted. Source code and executables for decompression programs must be provided to the appropriate PDS Discipline Node at the time an archive is delivered. Well documented decompression algorithms must be included in the archive

h. The compression and decompression software must be validated on a number of test data files to verify that the input and output files are identical. Thereafter, a random sampling of data products in the archive should be decompressed as part of the validation process.

i. The compressed products must be validated to comply with the specification or standard defining the compression algorithm used.

j. The compressed products, decompression algorithms, and decompression software must all be available for use by the PDS and its users on a royalty and license fee free basis

3. Compression of other files is allowed subject to the following conditions:

a. Lossless compression software from INFOZIP will be used; a PDS minimal label with pointer to INFOZIP will accompany the compressed file

b. PDS will capture the INFOZIP software tree at least annually and make it available for distribution

c. Files critical to understanding structure and basic content of the archive will NOT be compressed

Each Discipline Node accepting compressed data must keep an inventory of those holdings and take action to maintain the usability of the data as needed.

Amendment (2005-11-14) - add four words to 2c (passed):
NAIF– Acton Yes
Radio Science – Simpson Yes
Engineering Node – Crichton Yes
Atmospheres Node – Beebe Yes
Rings Node – Showalter Yes
Imaging Node – Gaddis Yes
GEO Node - Guinness Yes
PPI Node - Walker Yes
Small Bodies - Kolokolova Yes

Amendment (2013-11-19) – Added first line, restricting policy to PDS3
Passed by MC without opposition