Phoenix MECA Non-Imaging Bundles: Notes and Errata Regarding PDS3-to-PDS4 Migration

PDS Geosciences Node

June 2, 2026

Introduction

The Mars Phoenix Lander datasets were originally archived under NASA's Planetary Data System (PDS) PDS3 standard, and the PDS Nodes hosting the archives are actively migrating these archives to the PDS4 standard. Hereinafter, these notes apply specifically to the Phoenix Microscopy, Electrochemistry and Conductivity Analyzer (MECA) datasets archived at the PDS Geosciences Node.

The MECA instrument suite includes four instruments: the Atomic Force Microscope (AFM), Optical Microscope (OM), Thermal and Electrical Conductivity Probe (TECP), and Wet Chemistry Laboratory (WCL). The AFM, TECP, and WCL are known collectively as the MECA Non-Imaging Instruments. The Geosciences Node archives contain the data products from these instruments. The data from the OM are in separate archives (PDS3 data set IDs PHX-M-OM-2-EDR-V1.0, PHX-M-OM-3-RADIOMETRIC-SCI-V1.0).

Not all data products in the original MECA archives were PDS4-compliant, so these data files were reformatted (see "Reformatting" below). In general, the PDS3 archive volumes remain intact; no files have been removed. In addition to the reformatted data products, the only change is the addition of PDS4 labels and documentation files. In the data directories, the metadata in the attached or detached PDS3 labels have been copied to PDS4 labels, so that each data product now has both a PDS3 and a PDS4 label.

MECA Raw Data Types

The MECA Non-Imaging raw data consist of a single file containing binary data preceded by an ASCII PDS3 label. The body of the file is a binary dataset consisting of one of sixteen format types as described in the Software Interface Specification. Types 0-3 contain data pertinent to an AFM measurement; the actual scan data is a table in type 2. Types 4-6 are data returned by common electronics functions, including power monitoring (type 5). Type 7 contains all the TECP data. Types 8-15 pertain to the WCL: Type 8 contains data from all the potentiometric WCL sensors such as pH and ion-selective electrodes; type 9 contains conductivity data; types 10-14 contains various voltammetries and potentiometries, all in a common format; and type 15 contains pressure and temperature data.

Data types 2, 5, and 7 contained packed data fields that were not compatible with PDS4. These files have been reformatted (see "Reformatting" below). *meca_edr_sis.pdf* in the MECA non-imaging raw document collection contains further details on these and the other data product types.

MECA Calibrated and Derived Data Types

The MECA Non-Imaging calibrated and derived data consist of single ASCII data files, accompanied by detached PDS3 and PDS4 labels. Data are broken into 12 different data types according to instrument sub-system as described in *meca_rdr_sis.pdf* in the MECA non-imaging calibrated and derived document collection.

The AFM sub-system has 3 ASCII data types associated with it: two data files and one text description file. The AFM Scan Data Records (SDR) are AFM scan data that have been converted from DN to physical units. AFM Scan derivative records (SDD) are line-by-line first order spatial derivatives of the SDRs, processed using the Savitzky-Golay filter method. AFM data are grouped by measurement day.

The TECP sub-system yields four data types. The data are the electrical conductivity, frost point, relative permittivity, and the thermal conductivity measurements converted from DN to physical units. TECP data are in a time-series, with a single data file for each data type per observation, where an observation is a related set of measurements spanning no more than a day (sol).

The WCL sub-system has five data types, one for each sensor type or measurement method. The data are generally raw data that have been converted to physical units and grouped in a time-series corresponding to a single observation, which typically refers to a segment of a two-day chemistry experiment (during calibration, for example, or after acid addition).

In addition to the 12 data types there is a collection of derived data from a science team member, containing results from the recalibration of the TECP relative humidity sensor.

Archive Organization

Instead of the datasets and archive volumes in PDS3, products in PDS4 are organized into collections and bundles. A collection is a set of related products, which may be data products, document products, browse products, miscellaneous products, etc. A bundle is a set of related collections.

Two PDS4 bundles have been defined, one for the data from the raw volume and one for the data from the calibrated/derived volume. The raw bundle has 4 data collections (one for each instrument, plus one for MECA common electronics), two document collections, and one miscellaneous collection. The calibrated/derived bundle has 4 data collections (one for each instrument, plus one for the derived tecp_rhs dataset), two document collections, one calibration collection, and one miscellaneous collection. The table below shows the correspondence between the PDS3 volumes and PDS4 bundles.

PDS4 Bundles and Corresponding PDS3 Volumes and Datasets

Archive	PDS4 Bundle	PDS3 Volumes	PDS3 Data Set ID
Phoenix MECA Non-Imaging Raw Data Bundle	urn:nasa:pds:phx_meca_raw	PHXMEC_0XXX	PHX-M-MECA-2-NIEDR-V1.0
Phoenix MECA Non-Imaging Calibrated and Derived Data Bundle	urn:nasa:pds:phx_meca	PHXMEC_1XXX	PHX-M-MECA-4-NIRDR-V1.0

A bundle is identified by a file named bundle_*.xml (where * is some character sequence) in the root directory of a volume; this PDS4 label describes the bundle and lists the collections that belong to it. A collection, in turn, is identified by a PDS4 label named collection_*.xml in a subdirectory. The file collection_*_inventory.csv is found in the same directory as the collection label and lists the products that belong to that collection. Each product also has its own PDS4 label named *.xml. The PDS4 label and the data file it describes have the same base file name with a different extension.

Every bundle, collection, and product in PDS4 has a Logical Identifier (LID) that is guaranteed to be unique throughout the PDS. The LID is defined in the PDS4 label using the <logical_identifier> tag. For data products, the LID is analogous to PRODUCT_ID in a PDS3 label.

Reformatting

The original EDR type 2, 5, and 7 data files contained packed data fields, which were not PDS4-compatible. For the PDS4 archive, the bits in these fields were unpacked by adding 4 bits of zero to bring each field to a whole number of bytes. The values in each field were extracted using a big endian convention.

EDR Data Type 2: The PDS3 "DIR_CHANNEL_MASKS" column was initially two 4-bit fields packed together into 1 byte. These two fields were unpacked by adding 4 bits of zero *before* each of the 4-bit fields of data to ensure the actual data values remained unchanged. Together, the two fields of unpacked data total 2 bytes. For example, if the initial byte of data for the two 4-bit fields was 00100101, then the unpacked data became **0000**0010**0000**0101 (2 bytes total).

EDR Data Type 5: The PDS3 "POWER DATA" column was initially one single 12-byte field but has been split into 8 fields in a group for the PDS4 migration. Each of these 8 fields was initially 12 bits in the PDS3 dataset, and the fields were initially packed together so that they extended across byte boundaries. Four bits of zero were added to each of these fields *before* the 12 bits of data to ensure the actual data values remained unchanged. This resulted in a total of 8*4 bits (i.e., 4 bytes) being added to each record. For example, if the initial 12 bits of data comprising one field were 011010100010, the unpacked data for that field became **0000**011010100010 (2 bytes total).

EDR Data Type 7: The PDS3 "A TO D COUNTS" column was initially one single 12-byte field but has been split into 8 fields in a group for the PDS4 migration. Each of these fields was initially 12 bits in the PDS3 dataset, and the fields were initially packed together so that they extended across byte boundaries. Four bits of zero were added to each of these fields *before* the 12 bits of data to ensure the actual data values remained unchanged. This resulted in a total of 8*4 bits (i.e., 4 bytes) being added to each record. For example, if the initial 12 bits of data comprising one field were 011010100010, the unpacked data for that field became **0000**011010100010 (2 bytes total).

The table below lists the fields that were padded for each data type. *meca_edr_sis.pdf* in the MECA non-imaging raw document collection describes the original data product formatting.

EDR Data Type	PDS3 Column	PDS4 Fields	PDS3 Bytes	PDS4 Bytes
2	"DIR_CHANNEL_MASKS" (Column 11)	"DIRECTION_MASK", "CHANNEL MASK" (Fields 11-12)	1	2
5	"POWER DATA" (Column 13)	Group 1: "5V LOGIC VOLTS", "5V LOAD VOLTS", "15V LOAD VOLTS", "15V AFM VOLTS", "5V LOGIC CURRENT", "5V LOAD CURRENT", "15V LOAD CURRENT", "15V AFM CURRENT" (Fields 1-8)	12	16
7	"A TO D COUNTS" (Column 13)	"TECP_SAMPLE" (Group 1): "THERMOCOUPLE 1", "THERMOCOUPLE 2", "THERMOCOUPLE 3", "HUMIDITY", "ELECTRICAL CONDUCTIVITY", "BOARD TEMPERATURE", "DIELECTRIC", "HEATER CURRENT" (Fields 1-8)	12	16

The unpacking led to incorrect values in the “DATA LENGTH” and “SAMPLE SIZE” fields in the original EM5 and EM7 data files. These fields have been corrected.

Notes

1. MECA Raw Bundle
    a. Only the binary data were changed as part of the reformatting described above. No changes were made to the attached PDS3 label files.
    b. The attached PDS3 labels for the 3 reformatted EDR types no longer correctly describe the data files. Only the PDS4 labels should be used in these cases.
    c. Errors in PDS3 labels (in BYTES and ITEMS) for several EDR data types were corrected in the new PDS4 labels. No attempt was made to correct these errors in the PDS3 labels.
2. MECA Derived Bundle
    a. Several tables contained extra commas at the end of the records. These commas were removed.
    b. Sixteen AFM files had formatting issues in AFM_HEADER_TABLE. The .tab files were padded with spaces to make the record lengths consistent.
    sol046/af/fs046sdd_001_4e01150c0000a0.tab
    sol046/af/fs046sdd_002_4e01150c0000a0.tab
    sol046/af/fs046sdr_001_5a01150c0000a0.tab
    sol046/af/fs046sdr_002_5a01150c0000a0.tab
    sol064/af/fs064sdd_003_4e0117410000a0.tab
    sol064/af/fs064sdr_003_5a0117410000a0.tab
    sol068/af/fs068sdd_001_4e0101150000a0.tab
    sol068/af/fs068sdd_002_4e0101150000a0.tab
    sol068/af/fs068sdd_003_4e0101150000a0.tab
    sol068/af/fs068sdr_001_5a0101150000a0.tab
    sol068/af/fs068sdr_002_5a0101150000a0.tab
    sol068/af/fs068sdr_003_5a0101150000a0.tab
    sol099/af/fs099sdd_001_4e011b860000a0.tab
    sol099/af/fs099sdr_001_5a011b860000a0.tab
    sol132/af/fs132sdd_001_4e011f000000a0.tab
    sol132/af/fs132sdr_001_5a011f000000a0.tab