

Vespa Implementation Workshop 2021 PDS / PPI Overview

Steven Joy – Science Lead (UCLA)
In Sook Moon – UCLA Technical Lead
Chris Piker – U. of Iowa Technical Lead

Data Science Goals

- The Planetary Plasma Interactions (PPI) Node of the NASA Planetary Data System (PDS) is the primary archive for fields and particles (F&P) datasets from NASA planetary missions.
 - F&P data are used to study the transfer of mass, energy, and momentum from the Sun to the planets and other solar system bodies via the solar wind.
- The primary goal of the PPI Node in implementing a VESPA/epnTAP server is to improve to overall FAIRness of the PPI data archive
 - Currently the PPI server is completely open so planetary scientists anywhere in the world can access the system and download data
 - Data at the PPI Node can be difficult to work with (format issues) and there aren't as many tools to display and analyze the data as many users would like
 - Our hope is that making the data available to the user community through VESPA via an epnTAP server/service will improve the findability, interoperability, and reusability of the data in our archive – both at the data collection and individual data product levels
 - VESPA provides another mechanism for users to find our data
 - epnTAP should improve the interoperability of our archive
 - Europlanet and VO display and analysis tools should make our data easier to use thereby improving the reusability of our data.
 - PPI is also trying to improve the usability of some of its data with complicated data structures by reformatting the data into CDF formatted files, which are supported by many tools

Data Content and Format

- PPI selected five (5) data collections to be used in this workshop, with the goal of being able to access these collections via VESPA by the end or very shortly thereafter.
- Datasets that support Mercury and Mars planetary science, while also spanning a reasonable portion of the range of formats within PPI were selected
 1. Selected data collections from the [MESSENGER calibrated magnetometer bundle](#)
 2. The MESSENGER [KT17 magnetic field model residuals bundle](#)
 3. The MESSENGER [Mercury crustal magnetic field map bundle](#)
 4. The [MAVEN Mars Insitu Key Parameters data bundle](#)
 5. The [MAVEN Mars SWIA \(Solar Wind Ion Analyzer\) calibrated data bundle](#)
- All of these data are described by PDS4 metadata. Some of the bundles have multiple data collections, but we will focus on just one collection per bundle during the workshop. PPI will provide access to the remaining collections shortly thereafter.
- Data products in Items 1-4 are stored as ASCII fixed-width tables, while Item 5 are collections of CDF files containing multi-dimensional data.

PDS4 Collection Level Metadata

- PDS4 metadata are stored in XML format. Archives are organized into bundles (collections of collection), collections (groups of similar products), and individual products (file(s) plus label).
- Collection level metadata contain:
 - <logical_identifier> and <version_id> - The combination of LID & VID uniquely describe the collection
 - <title> Effectively the name of the collection
 - <Citation_Information> - Information about the authors, publication date, and collection **DOI**
 - <Modification_History> - Useful for users but not necessarily for searches
 - <Time_Coordinates> - Collection start/stop times
 - <Primary_Result_Summary> - Purpose (science, calibration, etc.) and processing level (calibrated, etc.)
 - <Observing_System> Spacecraft and Instrument names and documentation references
 - <Target_Identification> Target name and type (Mars, planet)
 - <Reference_List> - logical identifiers for internal (in PDS) and external (literature) references
 - <File_Area_Inventory> - CSV file containing LIDs & VIDs for every product in the collection
 - <records> - number of products in the collection

PDS4 Metadata for ASCII Data Tables

- Top level info is the same as for the collection level but it only applies to the individual products
 - LID & VID, title, citation, modification, time coordinates, primary results, target, observing system, references
- File level metadata
 - File name, creation date, size (bytes), records, md5_checksum
- File Structure metadata (# of records)
 - Header description (if any, # of records, header type)
 - Record level metadata (# of fields, groups, record length)
 - Tables are allowed to contain groups (repetitions) of fields (i.e. energy channels)
- Field Level Metadata
 - <name> - column name
 - <field number> column number, starting at 1
 - <field_location> byte offset to start of data column, starting at 1
 - <data_type> Integer, Real, Complex, various Date/Time formats, etc.
 - <field_length> size of field in bytes
 - <field_format> FORTRAN format of field (not required)
 - <description> Description of the column contents

PDS4 Metadata for CDF files

- Top level info is the same as for the collection level but it only applies to the individual products
 - LID & VID, title, citation, modification, time coordinates, primary results, target, observing system, references
- File level metadata
 - File name, creation date, size (bytes), records, md5_checksum
- File Structure metadata (# of records)
 - Header description (CDF header type, header bytes)
 - File contains data stored as arrays of various dimensions (<ARRAY> class)
 - All data (even constants) are described using the <ARRAY> class
 - Constants are described as degenerative arrays 1 axis, 1 element
 - Time and other scalars are stored as 1-D arrays (1 x n)
 - Data arrays are multi-dimensional arrays (2-D, 4-D)
 - <ARRAY> classes contain <name>, <offset>, <axes>, <axes_index_order>, <description> <data_type> and <Axis_Array> classes that describe the axes contents
 - <Axis_Array> classes contain <axis_name>, <elements>, <sequence_number>

1) MESSENGER MAG MBF Coordinates Data Collection

- <https://pds-ppi.igpp.ucla.edu/search/view/?f=yes&id=pds://PPI/mess-mag-calibrated/data/mbf>

Mercury body-fixed (MBF) coordinates, spacecraft position and B-field

- File Structure:

- No header, 12 fields, no groups, record length = 115 bytes, CR/LF line termination

Field Num	Name	Location	DataType	Length	Description
1	Year	1	ASCII_Integer	4	UTC Year
2	DayOfYear	6	ASCII_Integer	3	UTC day of year (1 = Jan 1)
3	Hour	10	ASCII_Integer	2	UTC hour
4	Minute	13	ASCII_Integer	2	UTC minute
5	Second	16	ASCII_Real (f6.3)	6	UTC second
6	TimeTagMET	23	ASCII_Real (f13.3)	13	Mission Elapsed Time (seconds) ~ SCLK
7	SC_X_MBF	37	ASCII_Real (f14.3)	14	Spacecraft MBF X position in km
8	SC_Y_MBF	52	ASCII_Real (f14.3)	14	Spacecraft MBF Y position in km
9	SC_Z_MBF	67	ASCII_Real (f14.3)	14	Spacecraft MBF X position in km
10	Bx_MBF	82	ASCII_Real (f10.3)	10	Bx MBF in nT
11	By_MBF	93	ASCII_Real (f10.3)	10	By MBF in nT
12	Bz_MBF	104	ASCII_Real (f10.3)	10	Bz MBF in nT

2) KT17 Model Residuals MBF Coordinates Data Collection

- <https://pds-ppi.igpp.ucla.edu/search/view/?f=yes&id=pds://PPI/mess-mag-kt17-model-residuals/data-sci-mbf>

KT17 Mercury Magnetic Field Model residuals - MBF coordinates

- File Structure:
 - No header, 20 fields, no groups, record length = 240 bytes, CR/LF line termination

Field Num	Name	Location	DataType	Len	Description
1	DateTimeUTC	1	ASCII_Date_Time_DOY	21	Record date/time (YYYY-DDDThh:mm:ss.sss)
2	TimeTagMET	23	ASCII_Real (f13.3)	23	Mission Elapsed Time (seconds) ~ SCLK
3	NAVG	37	ASCII_Integer	6	Number of samples in average
4	SC_X_MBF	44	ASCII_Real (f14.3)	14	Spacecraft MBF X position in km
5	SC_Y_MBF	59	ASCII_Real (f14.3)	14	Spacecraft MBF Y position in km
6	SC_Z_MBF	74	ASCII_Real (f14.3)	14	Spacecraft MBF X position in km
7	Bx_MBF	89	ASCII_Real (f10.3)	10	Average Bx MBF in nT
8	By_MBF	100	ASCII_Real (f10.3)	10	Average By MBF in nT
9	Bz_MBF	111	ASCII_Real (f10.3)	10	Average Bz MBF in nT
10	dBx_MBF	122	ASCII_Real (f10.3)	10	Residual Bx relative to KT17 model in nT
11	dBy_MBF	133	ASCII_Real (f10.3)	10	Residual By relative to KT17 model in nT
12	dBz_MBF	144	ASCII_Real (f10.3)	10	Residual Bz relative to KT17 model in nT

2) KT17 Model Residuals MBF Coordinates (con't)

- KT17 Mercury Magnetic Field Model residuals - MBF coordinates

Field Num	Name	Location	DataType	Len	Description
13	BxMI_MBF	155	ASCII_Real (f10.3)	10	KT17 Internal field model Bx MBF in nT
14	ByMI_MBF	166	ASCII_Real (f10.3)	10	KT17 Internal field model By MBF in nT
15	BzMI_MBF	177	ASCII_Real (f10.3)	10	KT17 Internal field model Bz MBF in nT
16	BxME_MBF	188	ASCII_Real (f10.3)	10	KT17 External field model Bx MBF in nT
17	ByME_MBF	199	ASCII_Real (f10.3)	10	KT17 External field model By MBF in nT
18	BzME_MBF	210	ASCII_Real (f10.3)	10	KT17 External field model Bz MBF in nT
19	Rhel_AU	221	ASCII_Real (f10.8)	12	Heliocentric range of Mercury in AU
20	ACTIDX	232	ASCII_Real (f5.1)	5	Activity index of Mercury's magnetosphere as percentage of maximum observed during the MESSENGER mission.

3) MESSENGER MAG Crustal Field Map Data Collection

- https://pds-ppi.igpp.ucla.edu/search/view/?f=yes&id=pds://PPI/mess-mag-crustal-field-map/data/Bm_output_0E360E&o=1

The product is a table showing vector field components at a constant altitude of 40 km extending from 35N to 75N latitude and covering all longitudes on a 0.5 deg x 0.25 deg grid.

- File Structure:
 - No header, 7 fields, no groups, record length = 68 bytes, CR/LF line termination

Field Num	Name	Location	Data Type	Length	Description
1	Longitude	1	ASCII_Real (f8.3)	8	Mercury Planetocentric (east) Longitude in degrees
2	Latitude	14	ASCII_Real (f7.3)	7	Mercury Planetocentric Latitude in degrees
3	B_total	22	ASCII_Real (f8.4)	8	Btotal – field magnitude in nT
4	B_radial	31	ASCII_Real (f8.4)	8	Outward radial field component in nT
5	B_east	40	ASCII_Real (f8.4)	8	Eastward field component in nT
6	B_north	49	ASCII_Real (f8.4)	8	Northward field component in nT
7	Altitude	58	ASCII_Real (f8.3)	8	Model altitude in km

4) MAVEN In Situ Key Parameters Data Collection

- <https://pds-ppi.igpp.ucla.edu/search/view/?f=yes&id=pds://PPI/maven.insitu.calibrated/data>
Key parameters data for the MAVEN in situ instruments (NGIMS, EUV, LPW, MAG, SEP, STATIC, SWEA, and SWIA)
- File Structure:
 - Multi-line (346) header, 235 fields, no groups, record length = 3765 bytes, CR/LF line termination
 - Too complex to describe easily, open example label and data file:

https://pds-ppi.igpp.ucla.edu/search/view/?f=yes&id=pds://PPI/maven.insitu.calibrated/data/2016/03/mvn_kp_insitu_20160306_v17_r04&o=1

5) MAVEN SWIA Calibrated Coarse Survey 3D Data Collection

- https://pds-ppi.igpp.ucla.edu/search/view/?f=yes&id=pds://PPI/maven.swia.calibrated/data/coarse_svy_3d

SWIA coarse survey 3d files with time-ordered fully calibrated ion distributions in units of differential energy flux derived from the SWIA Coarse distribution Survey telemetry

- **File Structure: CDF**

- CDF header
- Too complex to describe easily, open example label and data file:

https://pds-ppi.igpp.ucla.edu/search/view/?f=yes&id=pds://PPI/maven.swia.calibrated/data/coarse_svy_3d/2016/03/mvn_swi_l2_coarsesvy3d_20160304_v01_r01&o=1

Includes 1-D (scalars like time), 2-D, 4-D, and degenerative Arrays (1 axis, 1 element to describe constants like the geometric factor).