Update on Space Physics Data Facility (SPDF) Data Archives and Services

Robert Candey[1], Dieter Bilitza[1], Reine Chimiak[1], John Cooper[1], Leonard Garcia[1], Bernard Harris[1], Lan Jian[1], Rita Johnson[1], Tamara Kovalick[1], Nand Lal[1], Howard Leckner[1], Michael Liu[1], Sonya Lyatsky[1], Robert McGuire[1], Natalia Papitashvili[1], Uthra Rao[1], D Aaron Roberts[1], and Ronald Yurow[2]

[1]NASA Goddard Space Flight Center
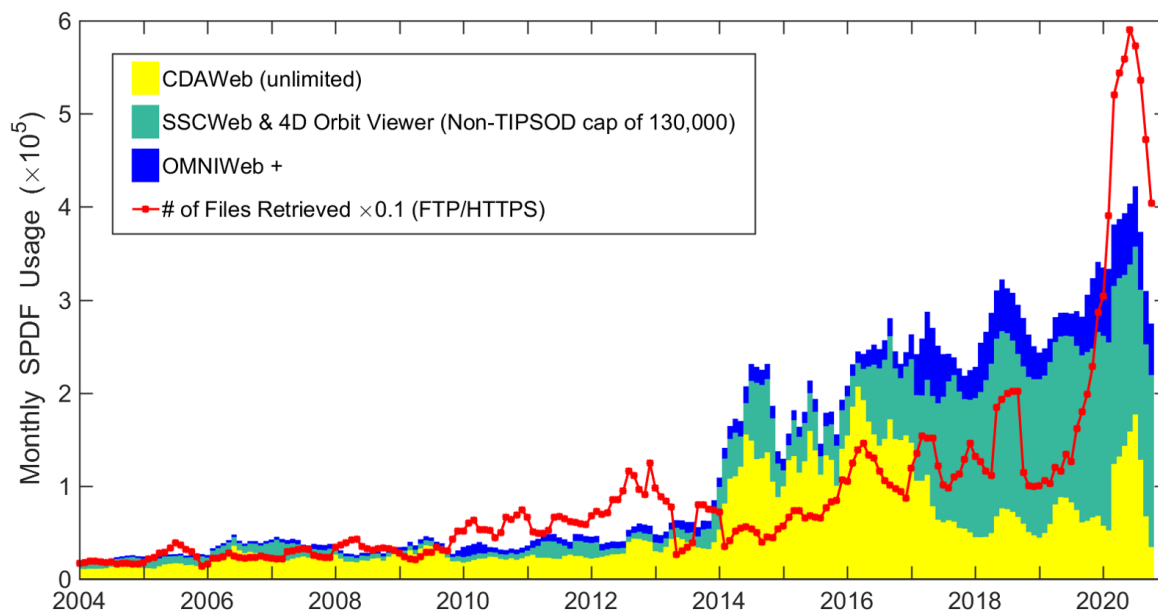[2]ADNET Systems Inc.

November 26, 2022

Abstract

The Space Physics Data Facility (SPDF https://spdf.gsfc.nasa.gov), as the non-solar NASA Heliophysics active final archive, works with current operating missions and the heliophysics community to ingest, preserve and serve a wide range of past and current public science-quality data from the mesosphere into the furthest reach of deep-space exploration. SPDF facilitates scientific analysis of multi-instrument and multi-mission datasets to enhance the science return of the many missions. SPDF develops and maintains the Common Data Format (CDF) and the associated ISTP/SPDF metadata guidelines. SPDF services include CDAWeb, which supports both survey and burst mode data with graphics, listings and data superset/subset functions. All public data held by SPDF are also available for direct file download by HTTPS or FTPS links from the SPDF home page (https://spdf.gsfc.nasa.gov). SPDF is currently receiving and serving from missions including Parker Solar Probe, MMS, Van Allen Probes, THEMIS/ARTEMIS, GOLD, ACE, Cluster, IBEX, Voyager, Geotail, Wind and many others, and >120 Ground-Based investigations. SPDF recently added support for ARASE/ERG and MAVEN as supplementary access at the requests of those missions, and is expecting Solar Orbiter and ICON data. SPDF also operates the multi-mission orbit displays and query services of SSCWeb and the Java-based 4D Orbit Viewer, as well as the Heliophysics Data Portal (HDP) discipline-wide data inventory and access service, and OMNIWeb and COHOWeb for near-Earth and deep-space solar wind plasma, magnetic field, and energetic particle database, respectively.

Update on Space Physics Data Facility (SPDF) Data Archives and Services

Robert Candey and the rest of the SPDF team

NASA Goddard Space Flight Center,　George Mason University,　Catholic University of America, ADNET, and Wyle

PRESENTED AT:

NASA HELIOPHYSICS ACTIVE FINAL ARCHIVE FOR NON-SOLAR DATA

SPDF <https://spdf.gsfc.nasa.gov> is the active and final archive of in-situ data from NASA heliophysics missions, including collaboration missions with other US and foreign agencies

- We also archive other data relevant to NASA heliophysics science objectives
  - Related data from planetary missions (e.g., MESSENGER, MAVEN, New Horizons)
  - Heliophysics data from some NOAA and DoD satellites (e.g., GOES, DSCOVR)
  - Ground-based magnetometers, aurora cameras, radars, etc., which are funded by NSF or other agencies/programs
- The data covers the space from the Sun to the local interstellar medium, including magnetosphere, ionosphere, thermosphere, and mesosphere (M-ITM) of the Earth and other applicable planets
- 34% of heliophysics papers in AGU journals acknowledged SPDF services and data
- SPDF provides three main science-enabling services besides archiving data
  - CDAWeb (Coordinated Data Analysis Web): browse, correlate, and display
  - SSCWeb (Satellite Situation Center): orbit/ground track displays and queries
  - OMNIWeb Plus: solar wind conditions, especially at bowshock nose
- SPDF enables multi-instrument, multi-mission heliophysics science
  - Specific mission/instrument data in context of other missions/data
  - Specific mission/instrument data as enriching context for other data
  - Ancillary services & software (orbits, data standards, special products)
- SPDF also builds critical infrastructures for the heliophysics data environment:
  - Common Data Format (CDF) https://cdf.gsfc.nasa.gov
  - Heliophysics Data Portal https://heliophysicsdata.gsfc.nasa.gov

SPDF data and services are widely used, as seen in this chart:

Poster at https://spdf.gsfc.nasa.gov/pub/documents/SPDF/presentations/
(https://spdf.gsfc.nasa.gov/pub/documents/SPDF/presentations/)

SPDF ACTIVITIES IN PAST YEAR AND NEAR FUTURE

- Added Parker Solar Probe (85), AIM, GOLD, IBEX data (40), and Explorer-35

- Working on Solar Orbiter (74) and ICON data, final data from Van Allen Probes, TWINS.

- New mission requirements <https://spdf.gsfc.nasa.gov/guidelines/archive_newdata_reqt.html>

- Exploring changes to become part of NASA's Heliophysics Digital Resource Library, such as additional outreach materials, user support, and support for data in cloud services

- Will add webservices for event lists for burst mode data and science events (CMEs, bow shock crossings, etc.) and support to SSCweb and CDAWeb

- Adding SPASE Resource ID and DOI to CDAWeb metadata and display

- Redirected alternative server names, such as cdaweb.sci.gsfc.nasa.gov, and switched FTP to FTPS, so please update your codes and scripts

- CDAWeb added netCDF support, variable_purpose attribute, IBEX mapped images

- CDF: planning new features and tools, systematize ISTP metadata guidelines, explore cloud support

- Adding support for end dates in filenames
  - <dataset>_<beginTime>-<endTime>_<varyingInfo>_v<version>.cdf   [hyphen or underscore between times]

- Developing alternatives to the Java-based 4D Orbit Viewer and SKTeditor tools
  - New SKTeditor may also enable the user to define the SPASE metadata at the same time as defining the internal metadata and structure of the CDF/netCDF to be created.
  - Developing prototypes of HTML5/Javascript-based interfaces for the 4D orbit viewer. We plan to expand this project to add interactive data plotting and sonification tied to the orbit display, similar to CDAWeb and perhaps data glyphs along the orbits as well.

SPDF FULLY SUPPORTS ACTIVE AND OLDER MISSIONS WITH MULTIPLE SERVICES AND ACCESS METHODS

- Direct file downloads via FTPS and HTTPS https://spdf.gsfc.nasa.gov/pub/data/

- Orbit and ground track displays/queries via SSCWeb and 4D Orbit Viewer, along with all other supported heliophysics missions

- CDAWeb services:
  - Plots and listings (ASCII, CSV, JSON)
  - Supersets or subsets by time & selected variables
  - Time-binning of data where appropriate
  - Web service interfaces (REST, SOAP, IDL, Matlab, Java, Python) https://cdaweb.gsfc.nasa.gov/WebServices/
  - New HAPI (Heliophysics API) https://cdaweb.gsfc.nasa.gov/hapi
  - Autoplot autoplot.org/help#CDAWeb
  - CDAWeb REST example (CDF fastest but other formats also: text, csv, json, png, gif, ps, pdf, nc, audio)
  - For instance, to get a CDF file containing the psp_fld_l2_mag_RTN data from the PSP_FLD_L2_MAG_RTN dataset for a time range:
    - https://cdaweb.gsfc.nasa.gov/WS/cdasr/1/dataviews/sp_phys/datasets/PSP_FLD_L2_MAG_RTN/data/20190307T100000Z,20190317T230000Z/psp_fld_l2_mag_RTN?format=cdf

- SPDF complement the services of the mission and instrument teams

- SPDF auto-ingest scripts check all supported mission data sites daily to retrieve new data files, and CDF files are validated and ingested

- Master CDFs add or improve metadata for use in CDAWeb

- The SPASE (Space Physics Archive Search and Extract, http://www.spase-group.org/) team use the master CDFs to generate SPASE IDs and descriptions for all datasets, to add entries to the Heliophysics Data Portal, https://heliophysicsdata.gsfc.nasa.gov

CDF STATUS AND RECENT DEVELOPMENT

- Released CDF 3.8.0.1
- Continued CDF support and general development, plus added features
- Use FILLVAL to fill the missing data if it exists, instead of the pad value
- Added options to cdfconvert (remove dimension without DEPEND_*, use FILLVAL if exists in place of pad value)
- Use FORMAT attribute for data listing if it exists
- Update zlib package to V2.1.11, for code improvements and bug fixes
- More generalized CDF epoch data encoding and parsing functions, default encoded epoch data now ISO-8601 format
- CDF epoch data conversion to/from Unix time
- String typed variables padded with a single space and followed by NULs
- cdf2skt tool option to choose how to display variable's metadata and data

CDF Plans

- High-level functions to read variables or whole CDF into a map structure for IDL, Java, Perl and C#
- Improve Windows installer, autoconf/make build/install, Mave/Ant/Gradle installs
- Improve documentation, beginner's guides, add to Wikipedia CDF entry
- Standardize ISTP/IACG Metadata Guidelines with version control, etc.
- New SKTeditor in Javascript or Python, perhaps also SPASE metadata creation
- Looking into supporting CDFs in cloud object storage, perhaps Zarr like netCDF is exploring
- Define CDF MIME type and international standard
- Apache 2 license in place of current custom license
- Update CDFML and its corresponding JSON representation with cdf.xsd use more specific datatype (e.g., xs:dataTime, xs:integer, xs:float, etc.) instead of just xs:string

ABSTRACT

The Space Physics Data Facility (SPDF https://spdf.gsfc.nasa.gov), as the non-solar NASA Heliophysics active final archive, works with current operating missions and the heliophysics community to ingest, preserve and serve a wide range of past and current public science-quality data from the mesosphere into the furthest reach of deep-space exploration. SPDF facilitates scientific analysis of multi-instrument and multi-mission datasets to enhance the science return of the many missions. SPDF develops and maintains the Common Data Format (CDF) and the associated ISTP/SPDF metadata guidelines. SPDF services include CDAWeb, which supports both survey and burst mode data with graphics, listings and data superset/subset functions. All public data held by SPDF are also available for direct file download by HTTPS or FTPS links from the SPDF home page (https://spdf.gsfc.nasa.gov). SPDF is currently receiving and serving from missions including Parker Solar Probe, MMS, Van Allen Probes, THEMIS/ARTEMIS, GOLD, ACE, Cluster, IBEX, Voyager, Geotail, Wind and many others, and >120 Ground-Based investigations. SPDF recently added support for ARASE/ERG and MAVEN as supplementary access at the requests of those missions, and is expecting Solar Orbiter and ICON data. SPDF also operates the multi-mission orbit displays and query services of SSCWeb and the Java-based 4D Orbit Viewer, as well as the Heliophysics Data Portal (HDP) discipline-wide data inventory and access service, and OMNIWeb and COHOWeb for near-Earth and deep-space solar wind plasma, magnetic field, and energetic particle database, respectively.