SUBMISSION AGREEMENT
BETWEEN
THE SPACE WEATHER PREDICTION CENTER
AND
THE NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION
FOR DEEP SPACE CLIMATE OBSERVATORY (DSCOVR) SPACE WEATHER DATA

2015-06-15

Introduction

This document represents the agreement that the Space Weather Prediction Center (NWS>NCEP>SWPC) (the "Provider") and the National Centers for Environmental Information (NCEI) (the "Archive") have reached for submitting the Provider's data, Deep Space Climate Observatory (DSCOVR) Space Weather Data, to the Archive for long-term preservation. It represents a joint effort between the Provider and the Archive to accurately document the agreement and the expectations between the two groups.

In order to ensure that the quality and integrity of the archived data is not compromised, the Provider and the Archive agree to maintain this agreement with accurate and up-to-date information through the life of the data submission.

DSCOVR is an operational environmental satellite that will monitor the solar wind and interplanetary magnetic field at the sun-earth L1 Lagrangian point. The real-time data will provide SWPC with an approximate 45 minute forecast of hazardous space weather conditions at earth. DSCOVR launched on 2015-02-11, with an on-orbit lifetime through FY19. Datasets to be provided are raw observatory telemetry data, Level 1 data, processed (Level 2+) space weather products, and associated Metadata (file) and Support Items (documents, reports, etc). Total daily electronic volume is estimated to be on the order of 4 GB [TBD].

Contacts

Persons included in all communications regarding the data submission.

Provider Contacts

Point of Contact, SWPC/DSCOVR Program Manager
Doug Biesecker
NOAA/NWS/SWPC
Scientist (Federal)
303-497-4474
doug.biesecker@noaa.gov
Working hours

SWPC/DSCOVR Responsible Scientist
Alysha Reinard
NOAA/NWS/SWPC
CIRES Research Associate
303-497-4748
alysha.reinard@noaa.gov
Working hours

SWPC/DSCOVR Systems Engineer
Jeff Johnson
NOAA/NWS/SWPC
CIRES Professional Research Assistant
303-497-6260
jeff.m.johnson@noaa.gov
Working hours

SWPC/DSCOVR Systems Engineer
Michael Burek
NOAA/NWS/SWPC
CIRES Professional Research Assistant
303-497-5222
michael.burek@noaa.gov
Working hours

Archive Contacts

Data Acquisition, NCEI-CO Data Administrator
Dan Kowal
NGDC
Data Administrator (Federal)
303-497-6118
dan.kowal@noaa.gov
Working hours

NCEI-CO/DSCOVR Data Manager
Archive Scientist
Systems Engineer
William Rowland
NCEI-CO>SGB
CIRES Professional Research Assistant
303-497-5086
william.rowland@noaa.gov
Working hours

NCEI-CO/DSCOVR Program Manager
William Denig
NCEI-CO>SGB
STP Division Chief
303-497-6323
william.denig@noaa.gov
Working hours

Data Overview

The Deep Space Climate ObserVatoRy (DSCOVR) satellite is a NOAA operated asset to be located at the L1 point. The primary space weather instrument is the PlasMag sensor complement consisting of a magnetometer (MAG), which will measure the local vector magnetic field and, a Faraday Cup (FC), which will measure the solar wind proton/alpha bulk properties (wind speed, density and temperature). The PlasMag data will be used for monitoring solar wind conditions in order to provide forecasts and nowcasts to SWPC customers. The PlasMag data will also be made available to scientists for sensor cal/val and for research purposes. DSCOVR launched on 2015-02-11, and reached station at the first Lagrange point on June 07, 2015. It will achieve Final Operational Capability (FOC) August, 2015. The required operating lifetime is at least two years, but End Of Life (EOL) is anticipated to be December 2019. The Archive will plan on receiving data from IOC to EOL and archiving the data in accordance with Data Center policies.

Datasets:

1 - Real-time Observatory Data (also know as virtual channel 0 (VC0) data - VC0 data are raw telemetry acquired from numerous ground antenna sites located worldwide. SWPC will acquire and package the VCO data into time sequenced day-files each consisting of CCSDS packets within a NetCDF-4 wrapper. The VC0 data rate is 20 kbps with an effective data volume (file size) of approximately 260 MB/day (including the NetCDF wrapper). The file specification for the VC0 day file is described in the DSCOVR Space-To-Ground Interface Control Document.

2 - Stored Observatory Data (VC1) - VC1 data are raw telemetry acquired by NASA. VC1 data will be acquired for periods of up to 4 hours per day. SWPC will acquire and package the VC1 data into time sequenced day files each consisting of the CCSDS packets within a NetCDF-4 wrapper. The VC1 data rate is 20 kbps with an effective data volume (file size) of approximately 40 MB/day (including the NetCDF wrapper). The VC1 data will be acquired sporadically by DSCOVR program (as requested by SWPC). The file specification for the VC1 day file is is described in the DSCOVR Space-To-Ground Interface Control Document.

3 - Solar Wind Data Products - The Solar Wind Data Products consist of Level 1 and Level 2+ products created by SWPC from the VC0 and VC1 raw telemetry. Space Weather Data Products are provided as day files in NetCDF-4 format. Volume is unknown at this time but estimated to be of order 3.5 GB/day. The file specification for the Space Weather Data Products day file is the DSCOVR SWPC ICD [TBR].

Note: a) The Archive is in discussions with the Provider regarding characteristics of the NetCDF Classic format insofar as ensuring compliance to Attribute Convention for Dataset Discovery (ACDD) conventions and completeness of global attribute information within each file. b) Level 1 data are unpacked, reformatted, and re-sampled Level 0 data with all calibration data and all supplemental information to be used in subsequent processing appended. Radiometric and geometric correction applied to produce parameters in physical units. Data generally presented in full space/time resolution. (Level 1 data definition per the GSRD)

4 - Metadata and Support Items - Periodically SWPC will provide NGDC with relevant ancillary data in the form of Metadata and Support Items. Ancillary data includes calibration tables, test data, instrument and satellite status records, documents and other relevant information. The file specifications for these items are in the locations discussed in section 3.2.1.4.4. The only Support Items currently identified are the Telemetry Database, Faraday Cup Calibration files, and Magnetometer Calibration files. Metadata data volumes are are expected to be less than 100KB/day. The Telemetry database is approximately 5MB in size, but will only change 2-3 times over the lifetime of the mission.

Applicable and Reference Documents

Documents applicable to or referenced from this agreement.

1. - DSCOVR L1 Requirements Document (L1RD), Version 1.1, 25 Oct 2013)
2. - DSCOVR Project Ground Segment Requirements Document (GSRD), Rev -, 08 Feb 2012
3. - DSCOVR Request To Archive, 08 Feb 2013
4. - DSCOVR Space-To-Ground Interface Control Document
5. - SWPC-NGDC Interconnection Security Agreement, dated Sept 2012
6. - DSCOVR/NGDC Interface Control Document (ICD) [DSCOVR-ICD-002042]
7. - NESDIS Quality Procedure [NQP]-4100; NOAA/NESDIS Space Flight Program Management Requirements for Assisted Acquisitions (draft: 2/2013)
8. Multi-Mission Flight Dynamics Operational Agreement (MMFD-OA); DSCOVR Program
9. DSCOVR Space Weather Prediction Center Interface Control Document (DSCOVR SWPC-ICD); DSCOVR Program
10. DSCOVR Database Management Plan (DSCOVR-PLAN-001755); DSCOVR Program

Submission Scope

Active Submission Period

2015-06-01 - 2019-12-31

Data Types

Below is a summary of the data sizing and submission schedule by data type group. Enter information on at least one data type.

Data Type Name	Data Sizing	Submission Schedule
Real-Time Observatory Telemetry (VC0 day files)	260 MB [TBD]	Once per day after 00 UTC
Stored Observatory Telemetry (VC1 day files)	40 MB [TBD]	Once per day after 00 UTC
Solar Wind Data Products	3.5 GB [TBD]	Once per day after 00 UTC
DSCOVR/SWPC Metadata	100KB	Once per two weeks after 00 UTC, when applicable. Other transfer options for Supplemental Data may be used
Support Items	5MB	Approximately 3x during DSCOVR lifetime

Reviews and Testing

The DSCOVR program development is compliant with the NESDIS Quality Procedure [NQP]-4100 document for NOAA/NESDIS Space Flight Program Management Requirements for Assisted Acquisitions. The end-to-end data flows from the observatory to users are within the purview of the DSCOVR Ground System (GS) involving, as needed; 1) Data Format Reviews including compliance and completeness assurance, 2) Systems Engineering Reviews relevant to the overall data flows within the DSCOVR GS, 3) Connectivity Testing of the SWPC-NGDC interface for ingesting Observatory products and 4) Access Testing involving ingest test data into NGDC systems and making targeted data searchable and accessible through supported online tools.

Providing System

Identification of the system providing the data to NCEI.

System Name:	Space Weather Prediction Center
System Owner:	NOAA/NWS/SWPC
Physical Location:	Boulder, CO
Additional Information:	Reference: SWPC-NGDC Interconnection Security Agreement, dated Sept 2012

Transfer Interface

Reference the SWPC-NGDC Ingest ICD [TBD]. Secure File Transfer Protocol (SFTP) will be used to pull from SWPC server "dido-02". Details/Recommendations on the Transfer Interface will be provided by the NGDC Architecture Review Committee (ARC).

Submission File Inventory

Information on each submitted file type from the Provider. Information on multiple file types can be added below.

File Type Name: Real-Time Observatory Telemetry

File Name Pattern:

'<type>_<dataID>_<satellite>_<startDate>_<endDate>_<processDate>.<extension>.<compression>'

File Name Field Definitions:

<type>

"oe" - Operational Environment Data

"ot" - Operational Environment Test Data

"ie" - Integration and Test Environment Data

"it" - Integration and Test Test Data

<dataID>

"rot" - Real-Time Observatory Telemetry

<satellite>

"dscovr" - DSCOVR satellite

<startDate>

"smmddyyyyhhmmss" - data start time (UTC)

<endDate>

"emmddyyyyhhmmss" - data stop time (UTC)

<processDate>

"pmmddyyyyhhmmss" - data process time (UTC)

<extension>

"nc" - NetCDF extension

<compression>

"gz" - gzip

Example File Name:

oe_rot_dscovr_s01012015000000_e01012015235959_p01022015030000.nc.gz

File Format: netCDF

File Compression: gzip

File Size Average: 260MB

File Size Range: 1MB to 260MB

File Count (Rate): 1 file per day

Data Volume (Rate): 260 MB per day [TBD]

Submission Schedule: after 00 UTC, each day

Additional Information: CCSDS Telemetry (binary) encoded according to DSCOVR Telemetry standards, wrapped in netCDF.

Descriptive Information Attributes:

None: date/time fields are usually used as descriptive attributes for this file.

File Type Name: Stored Observatory Telemetry

File Name Pattern:

'<type>_<dataID>_<satellite>_<startDate>_<endDate>_<processDate>.<extension>.<compression>'

File Name Field Definitions:

<type>

"oe" - Operational Environment Data

"ot" - Operational Environment Test Data

"ie" - Integration and Test Environment Data

"it" - Integration and Test Test Data

<dataID>

"sot" - Stored Observatory Telemetry

<satellite>

"dscovr" - DSCOVR satellite

<startDate>

"smmddyyyyhhmmss" - data start time (UTC)

<endDate>

"emmddyyyyhhmmss" - data stop time (UTC)

<processDate>

"pmmddyyyyhhmmss" - data process time (UTC)

<extension>

"nc" - NetCDF extension

<compression>

"gz" - gzip

Example File Name:

oe_sot_dscovr_s01012015000000_e01012015235959_p01022015030000.nc.gz

File Format: netCDF

File Compression: gzip

File Size Average: 60MB

File Size Range: 1MB to 40MB

File Count (Rate): 1 file per day

Data Volume (Rate): 60 MB per day [TBD]

Submission Schedule: after 00 UTC, each day

Additional Information: CCSDS Telemetry (binary) encoded according to DSCOVR Telemetry standards, wrapped in netCDF.

Descriptive Information Attributes:

None: date/time fields are usually used as descriptive attributes for this file.

File Type Name: Solar Wind Data Products
File Name Pattern:
'<type>_<dataID>_<satellite>_<startDate>_<endDate>_<processDate>.<extension>.<compression>'
File Name Field Definitions:
<type>
"oe" - Operational Environment Data
"ot" - Operational Environment Test Data
"ie" - Integration and Test Environment Data
"it" - Integration and Test Test Data

<dataID>
"fc0" - Faraday Cup L0
"fc1" - Faraday Cup L1 (Full resolution)
"f3s" - Faraday Cup L1, 3s window (avg)
"f1m" - Faraday Cup L1, 1m window (avg)
"mg0" - Magnetometer L0
"mg1" - Magnetometer L1 (Full resolution)
"m1s" - Magnetometer 1s averages
"m1m" - Magnetometer 1m averages
"pop" - Predicted Orbit Product (s/c location)

<satellite>
"dscovr" - DSCOVR satellite

<startDate>
"smmddyyyyhhmmss" - data start time (UTC)

<endDate>
"emmddyyyyhhmmss" - data stop time (UTC)

<processDate>
"pmmddyyyyhhmmss" - data process time (UTC)

<extension>
"nc" - NetCDF extension

<compression>
"gz" - gzip
Example File Name:
oe_sw0_dscovr_s01012015000000_e01012015235959_p01022015030000.nc.gzip
File Format: netCDF
File Compression: gzip
File Size Average: 3.5GB
File Size Range: 1GB to 3.5GB
File Count (Rate): 1 file per day
Data Volume (Rate): 3.5 GB per day

Submission Schedule: after 00 UTC, each day

Additional Information: Files identified so far are netCDF-Classic.

Descriptive Information Attributes:

None: date/time fields are usually used as descriptive attributes for this file.

File Type Name: Metadata

File Name Pattern:

'<type>_<dataID>_<satellite>_<metadataDate>_<description>.<extension>.<compression>'

File Name Field Definitions:

<type>

"oe" - Operational Environment Data

"ot" - Operational Environment Test Data

"ie" - Integration and Test Environment Data

"it" - Integration and Test Test Data

<dataID>

"smd" - SWPC/DSCOVR Metadata

<satellite>

"dscovr" - DSCOVR satellite

<metadataDate>

"mmmddyyyyhhmmss" - date of delivery

<description>

variable length descriptive text

<extension>

"pdf" - Portable Document Format extension

"xlsx" - Excel extension or "xlsx"

"docx" - MS Word extension of "docx"

- Other electronic formats, as needed

<compression>

"gz" - gzipDefinitions of the fields in the file name pattern

Example File Name:

oe_smd_dscovr_s01012015000000_e01012015235959_p01022015030000.nc.gzip

File Format: [TBD]

File Compression: gzip

File Size Range: 3KB to 100KB

File Count (Rate): 3 files per 2 weeks

Data Volume (Rate): Less than 100KB/2 weeks

Submission Schedule: bi monthly

Additional Information: No specific metadata files have currently been identified.

Descriptive Information Attributes:

Attribute	Source	Use
metadata>>description	Description of metadata submission	Contains pertinent information required for sensor data stewardship

File Type Name: Support Items
File Name Pattern:
"<type>_<dataID>_<satellite>_<startDate>_<endDate>_<processDatae>.<extension>.<compression>"
File Name Field Definitions:
<type>
"oe" - Operational Environment Data
"ot" - Operational Environment Test Data
"ie" - Integration and Test Environment Data
"it" - Integration and Test Test Data

<dataID>
"tmd" - Telemetry Database
"fcc" - Faraday Cup Calibration (Plas)
"mgc" - Magnetometer Calibration (Mag)

<satellite>
"dscovr" - DSCOVR satellite

<startDate>
"sYYYYMMDDhhmmss" - data start time (UTC)

<endDate>
"eYYYYMMDDhhmmss" - data stop time (UTC)

<processDate>
"pYYYYMMDDhhmmss" - data process time (UTC)

<extension>
"nc" - NetCDF extension

<compression>
"gz" - gzip
Example File Name:
oe_tmd_dscovr_s20151229000000_e20151229235959_p20151229003000.nc.gz
File Format: netCDF
File Compression: gzip
File Size Average: 100KB
File Size Range: 100KB to 5.1MB
File Count (Rate): 1 5MB File, up to three times during DSCOVR Program plus 2 files totalling <100KB every 2 weeks
Data Volume (Rate): Episodic
Submission Schedule: Telemetry Database up to three times during DSCOVR Program, others as often as every 2 weeks
Additional Information: Calibration Data are expected to have a data volume less than 100KB/day. The Telemetry database is approximately 5MB in size, but will only change 2-3 times over the lifetime of the mission.

Descriptive Information Attributes:
None: date/time fields are usually used as descriptive attributes for this file.

Submission Manifest

A submission manifest file with a 32-character MD5 checksum value is required for each submitted file in order to ensure the integrity of the submitted data.

File Content Specification:
SWPC will provide Message-Digest Algorithm (MD5) checksum information with all files ingested by NGDC. Placing this checksum file, with an .md5 extension, into the directory /ngdc/archives/rtws/ from "dido-02" is the mechanism that SWPC uses to notify NGDC that a new SIP has been made available. NGDC will independently compute the checksum of the incoming SIP as a form of validation. If the checksums match, then after the data have been fully archived NGDC will write the NGDC checksum, with an extension of .rcp, to the directory "/ngdc/receipts/rtsw" on "dido-02".

File Transmission:
1 MD5 file per SIP per day

File Name Pattern:
"<SIP_filename_minus_extension>.<extension>"

File Name Definitions:
<SIP_filename_minus_extension>
<type>_<dataID>_<satellite>_<startDate>_<endDate>_<processDatae>
The SIP filenames are defined above. This file will use the same filename that was used for the incoming SIP, without the extensions.

Example File Name:
"oe_tmd_dscovr_s20151229000000_e20151229235959_p20151229003000.nc.gz.md5"

Archive Ingest

Ingest processing steps at the Archive and communication with the Provider.

Receipt Verification:
The Archive will use the provided file name and 32-character MD5 checksum value to verify the integrity of a delivered file.

Error Reconciliation:
The Archive will report any problems or errors with file integrity, file name, checksum validation, or other errors that inhibit the data ingest and archive to the Provider. A new corresponding submission manifest will be required for files re-submitted by the Provider.

Receipt Confirmation:
The Archive will provide an inventory of the data ingested once it is completed or as requested by the Provider.

Quality Assurance:
In the event of erroneous data the Provider will be requested to re-process and re-transmit the data. It is at the Archive's discretion to determine the final disposition of erroneous data.

Archive File Packaging:
Description of file packaging or re-naming by the Archive upon ingest.

Archive Storage

Archive attributes of each archived file type.

Archive File Type Name: Real-time Observatory Telemetry data
Archive File Attributes/IDs:

Attribute/ID Type	Value
Archive File ID	Archive filename and file types are identical to the original submission package.

Archive File Type Name: Stored Observatory Telemetry data
Archive File Attributes/IDs:

Attribute/ID Type	Value
Archive File ID	Archive filename and file types are identical to the original submission package.

Archive File Type Name: Solar Wind Product data
Archive File Attributes/IDs:

Attribute/ID Type	Value
Archive File ID	Archive filename and file types are identical to the original submission package.

Archive File Type Name: SWPC/DSCOVR Metadata
Archive File Attributes/IDs:

Attribute/ID Type	Value
Archive File ID	Archive filename and file types are identical to the original submission package.

Archive File Type Name: SWPC/DSCOVR Support Items
Archive File Attributes/IDs:

Attribute/ID Type	Value
Data Family ID	DSCOVR

Archive Updates

The Archive will archive the day-files as received from the Provider. In the event that erroneous data are discovered the Archive will request that the Provider reprocess and resubmit new day-files, as needed, to replace the original files and to ensure the quality of the archived dataset. Resubmitted day-files will be identified using the File Name Field Definition for the <processDate> at a minimum (see Submission File Inventory). Other File Name Field Definitions may be affected.

Retention Schedule

The data will be retained in the Archive for long-term preservation in accordance with NOAA data management standards. Information on data usage and archive value may be used for making decisions on continuing the duration of the archive. The disposition according to the NOAA Data Centers Record Schedule applies the following for disposition: retain Input/Master Data Files (L0/L1b) for not less than 75 years after the end of mission; retain Output Files (L2+) until no longer needed but not less than 3 years after end of mission. For the DSCOVR L2+ data the Archive agrees to retain the data for a period of not less than 10 years. Prior to deletion the Archive will solicit retention recommendations from the Provider and/or the heliophysics community.

(Notional) Disposition: 2095-12-30

Constraints

Constraint Type	Description
Access	External users will have unconstrained access to Operational Environment (OE) Solar Wind Data Products
Access	External users will be provided access to non-Operational Environment (OE) Solar Wind Data Products on an exception basis only.
Access	NGDC data steward will have unconstrained access to all Operational Environment (OE) and non-OE data.
Access	The Support Items object "Telemetry Database" is restricted. External users will be provided access on an exception basis only.
Access	External users will have unconstrained access to Real-Time Observatory Telemetry, Stored Observatory Telemetry, and DSCOVR/SWPC Metadata.
Access	All users of restricted datasets must obtain authorization via SWPC prior to obtaining access to the data via the NGDC Point of Contact.

User Community

1. The Operational Space Weather Community, in particular the Space Weather Prediction Center at NOAA and the instrument team at the NASA Goddard Space Flight Center (GSFC). The GSFC team, headed up by Dr Adam Szabo (adam.szabo-1@nasa.gov), is responsible for monitoring the general health and status of the PlasMag instrument, including sensor long-term calibration.

2. The heliophysics research community, including scientists at national laboratories and universities seeking to better understanding Space Weather phenomena and societal impacts. Representatives of this community include Professors Tamas Gambosi (tamas@umich.edu) from the University of Michigan and W. Jeffrey Hughes (hughes@bu.edu) from Boston University, both of whom are responsible for the development of distinct sun-to-earth space weather models.

User Documentation and Metadata

The Archive will be responsible for creating standard archival metadata for DSCOVR. The Provider will supply information to the Archive for writing and maintaining standard archive metadata, which includes data discovery information, references and data archive access links for users. The archival metadata will be created using the ISO 19115 standard.

Representation Information Items

This information is currently [TBD]. Detailed data formats will be derived from the DSCOVR Space-to-Ground ICD [TBD] and the DSCOVR/NGDC ICD [TBD]. General statement - For data to be useful to users, present and future, its format specification and characteristics must be documented and preserved with the data. Representation Information provides users with syntax (structure) and/or semantics (meaning) to decode the encoded data.

None

Preservation Descriptive Information Items

Background information about the DSCOVR Program, SWPC data processing, relevant papers, etc. will be included as these data become available. Metadata record and global attribute information in the NetCDF records will provide some of this information as well. General statement - Preservation Descriptive Information items contain context, provenance, and/or quality information for the data.

None

Access and Dissemination

Access and dissemination for unrestricted DSCOVR datasets will occur via a DSCOVR interface to the NGDC NEXT system, which was developed through NGDC's Agile development process. Stakeholder input was a key to success. The NEXT interface was validated after completion in Feb. 2014, and has been regression tested from the completion date through the present.

Additional Terms

Development and acceptance of the DSCOVR/NGDC ICD was a required element for the Archive support of the DSCOVR program.
NGDC/SWPC ICD contains the connection details of the interface used by NGDC to acquire DSCOVR data. This document is currently being developed.