

**SUBMISSION AGREEMENT
BETWEEN
THE NATIONAL CENTER FOR ENVIRONMENTAL INFORMATION
AND
THE NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION
FOR GOES-R LEVEL 0 SPACE WEATHER DATA**

2017-11-15

Introduction

This document represents the agreement that the National Center for Environmental Information (NOAA > NESDIS > NCEI) (the "Provider") and the National Centers for Environmental Information (NCEI) (the "Archive") have reached for submitting the Provider's data, GOES-R Level 0 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.

Note that NCEI is archiving L0 Space Weather data on its own initiative. Thus the SA agreement is between NCEI and itself.

Contacts

Persons included in all communications regarding the data submission.

Provider Contacts

Point of Contact, Data Manager	Involved in Level Zero Storage System (LZSS)
Margaret Tilton	Ananth Rao
DOC/NOAA/NESDIS/NCEI	GOES-R Ground Segment Project (NASA)
GOES-R Data Manager	Contractor
303-497-6223	(301) 286-4795
margaret.tilton@noaa.gov	ananth.rao-1@nasa.gov

Archive Contacts

Data Acquisition, Managing archive of L0 data
Margaret Tilton
DOC/NOAA/NESDIS/NCEI
GOES-R Data Manager
303-497-6223
margaret.tilton@noaa.gov

Data Overview

The data to be archived are Level 0 (L0) data from the GOES-R series satellites (GOES-R/T/U/V). These data will be archived in two formats:

1.) Full-resolution unformatted data received from the Level Zero Storage System (LZSS). These are instrument and spacecraft calibration data with communications artifacts (e.g. synchronization frames, communications headers) removed. These data will be tarred on a daily basis, and the resulting tarball will be the AIP. Additional information:

- Files come from each of the GOES-R space-looking instruments--SUVI, EXIS, MAG and SEISS.
- Science, engineering and diagnostic data are included, along with their instrument calibration parameters.
- Data files are CCSDS packets in a netCDF container.
- The temporal reference for L0 products is Coordinated Universal Time (UTC).

All higher-level GOES-R products are derived from this L0 data. The purpose of the L0 Products is to gather and trend geophysical data from the sun to the Earth, including Earth observations from space.

L0 data files are further described in the Product Definition and User's Guide (PUG) (DOC Control # 7035538, Rev F, Volume 2), as follows:

EXIS: Tables 4.2.1, 4.2.2-1, 4.2.2-2

MAG: Tables 4.4.1, 4.4.2-1, 4.4.2-2

SEISS: Tables 4.5.1, 4.5.2-1, 4.5.2-2

SUVI: Tables 4.6.1, 4.6.2-1, 4.6.2-2

2.) Aggregated and reformatted L0 data. STP plans to reformat L0 data for all instruments so that CCSDS packets have been decoded and are in user-accessible format. For all instruments except SUVI, STP will aggregate the reformatted files into daily .nc files. (SUVI data consist of individual images and are not easily aggregated.) Note that aggregated and reformatted data are not available yet, and they may not be so until 2019 or even later.

In addition, NCEI will be archiving L0 ISO-SERIES Metadata files for each Space Weather instrument. These change infrequently and their size is negligible.

Applicable and Reference Documents

Documents applicable to or referenced from this agreement.

1. Product Definition and User's Guide (PUG) (DOC Control # 7035538, Rev F, Volume 2)
2. Ground System Interfaces DTWT (Data Transfer Table) Modeling, December 5, 2016
3. Spacecraft Telemetry and Command (ST&C) Handbook (GOES-RQ-11-0159)
4. GOES-R Series Metadata Model, CDRL SE-21 (Harris DCN 8021130)
5. GOES-R Ground Segment Acronym List and Glossary (Harris DCN 7034325)
6. CORE GS NAMING CONVENTIONS, Sep, 2014 (Harris DCN 7043185)
7. GOES-R Level Zero Storage System (LZSS) User's Guide (416-R-LZSSUSG-0283)
8. GOES-R Ground Segment Remote Access (no doc number)
9. GOES-R Series Ground Segment (GS) Level Zero Storage Service (LZSS) Design Document (416-R-LZSSDD-0279)

Submission Scope

Active Submission Period

2016-02-01 - 2036-10-01

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
raw L0 EXIS file (all files tarred and gzipped)	80 MB per file	1 file per satellite per day
raw L0 SUVI file	360 MB per file (files are hourly)	24 files per satellite per day (one per hour)
raw L0 SEISS file	162 MB per file	1 file per satellite per day
raw L0 MAG file	63 MB per file	1 file per satellite per day
raw L0 ISO Series Metadata	Negligible (~210 MB for *all* L0 ISO metadata in the entire GOES-R series lifetime)	1 file per satellite per week
aggregated and reprocessed L0 EXIS .nc file	160 MB per file	1 file per satellite per day
reprocessed L0 SUVI .nc file (not aggregated)	480 MB per file (files are hourly)	1 file per satellite per day
aggregated and reprocessed L0 SEISS .nc file	324 MB per file	1 file per satellite per day
aggregated and reprocessed L0 MAG .nc file	126 MB per file	1 file per satellite per day

Reviews and Testing

There is md5 checksum validation performed to ensure data integrity.

Providing System

Identification of the system providing the data to NCEI.

System Name: Level-0 Storage System (LZSS)
 System Owner: Steve Grippando (Steven.Grippando@noaa.gov)
 Physical Location: Wallops Island LZSS server
 Additional Information: All files to be retrieved from LZSS are netCDF.

Transfer Interface

Transfer protocol: The LZSS is designed to work via a NOAA VPN connection, which requires a user with a CAC card. This system precludes true M2M access. Once a day, an NCEI user logs on to the LZSS system via CISCO AnyConnect software and retrieve data via scripts. Data will then be ingested into NCEI's archive via the Agile team's Common Ingest system. (Note the Common Ingest software is not working reliably for GOES-R space weather data. However, the Boulder Agile team has just incorporated new features that are expected to solve this issue.)

Bandwidth: There is finite bandwidth between the Ground Segment (LZSS) and the external networks. This is limited to 1gbps at WCDAS, but the ground system is looking into increasing this to 10 gbps. This bandwidth is shared between remote users and GOES-R external partners (including AWIPS/NWS). NCEI plans to open multiple VPN connections simultaneously so that data can be transferred at an adequate speed.

Submission File Inventory

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

File Type Name: Raw L0 instrument data, tarred and gzipped

File Name Pattern:

[env]_[inst]-L0_[platform]_s[start_date]_e[end_date]_v[version].tar.gz

File Name Field Definitions:

[env] = environment. Valid values: OR, OT, IR, IP, IS, IT

[inst] = instrument. Valid values: EXIS, SEIS, MAG, SUVI

[platform] = satellite. Valid values: G16, G17, G18, G19

[start_date] = start of observation interval in YYYYmmddHHMMSS format

[end_date] = end of observation interval in YYYYmmddHHMMSS format

[version] = version in X_Y_Z format.

Example File Name:

OR_EXIS-L0_G16_s20150826000000_e20150826235959_v0_1_1.tar.gz

File Format: netCDF-4

File Compression: gzip

File Size Range: 63MB to 360MB

File Count (Rate): 27 / day / satellite

Data Volume (Rate): 9 GB / satellite / day (for satellite in operational mode); 225 MB / satellite / day (for storage mode)

Submission Schedule: After launch until end of GOES-R Series

Additional Information: Raw files (i.e., files after untarring and decompressing) are simply CCSDS telemetry packets in a netCDF wrapper. A CCSDS decoder will be needed to convert files to human-readable form.

Descriptive Information Attributes:

Attribute	Source	Use
Instrument	File name	For search, results display, and/or cross-referencing
Start and end date	File name	For search, results display, and/or cross-referencing
Platform	File name	For search, results display, and/or cross-referencing
Environment	File name	For search, results display, and/or cross-referencing

File Type Name: Reformatted and aggregated L0 instrument data, tarred and gzipped

File Name Pattern:

[env]_[inst]-L0_[platform]_[start_date]_[end_date]_reformat_v[version].nc.tar.gz

File Name Field Definitions:

[env] = environment. Valid values: OR, OT, IR, IP, IS, IT

[inst] = instrument. Valid values: EXIS, SEIS, MAG, SUVI

[platform] = satellite. Valid values: G16, G17, G18, G19

[start_date] = start of observation interval in YYYYmmddHHMMSS format

[end_date] = end of observation interval in YYYYmmddHHMMSS format

[version] = processing version in X_Y_Z format

Example File Name:

OR_EXIS-L0_G16_s20150826000000_e20150826235959_reformat_v0_0_1.tar.gz

File Format: netCDF

File Compression: gzip

File Size Range: 126GB to 480MB

File Count (Rate): 27 / satellite / day

Data Volume (Rate): 35 GB / satellite / day (for satellite in operational mode); 450 MB/ satellite / day (for storage mode). NOTE: Since NCEI hasn't yet aggregated and reprocessed L0 data, these data sizes are estimated

Submission Schedule: After launch until end of GOES-R Series

Additional Information: These files don't exist yet, but NCEI instrument scientists hope to create them eventually. Some reformatting will have been performed, so that files are formatted by scientific variables.

Descriptive Information Attributes:

Attribute	Source	Use
Instrument	File Name	For search, results display, and/or cross-referencing
Start and End date	File Name	For search, results display, and/or cross-referencing
Platform	File Name	For search, results display, and/or cross-referencing
Environment	File Name	For search, results display, and/or cross-referencing
Version Number	File Name	For search, results display, and/or cross-referencing

File Type Name: Raw L0 ISO-SERIES metadata files, tarred and gzipped.

File Name Pattern:

[env]_[inst]-L0-ISO-SERIES_v[version].tar.gz

File Name Field Definitions:

[env] = environment. Valid values: OR, OT, IR, IP, IS, IT

[inst] = instrument. Valid values: EXIS, SEIS, MAG, SUVI

[version] = version in X_Y_Z format.

Example File Name:

OR_EXIS-L0-ISO-SERIES_v0.0.1.xml.tar.gz

File Format: XML

File Compression: gzip

File Size Average: 0.13MB

File Count (Rate): 4 files / week

Data Volume (Rate): 0.52 MB / week

Submission Schedule: After launch until end of GOES-R Series

Additional Information: Files will be weekly XML files that are tarred and gzipped.

ISO-SERIES files are for the entire GOES-R series collection and thus are not satellite-specific.

Descriptive Information Attributes:

Attribute	Source	Use
Instrument	File name	For search, results display, and/or cross-referencing
Environment	File name	For search, results display, and/or cross-referencing
Version Number	File name	For search, results display, and/or cross-referencing

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:

NCEI will supply manifests for each day's worth of data. Each manifest will list full file names and associated MD5 checksums for all files (raw, reformatted, and ISO-SERIES) submitted that day.

File Transmission:

MD5 files are required inputs for NCEI's Common Ingest system. These manifest files do not get archived, but Common Ingest uses them to verify the integrity of submitted data files.

File Name Pattern:

[FileName].md5

File Name Definitions:

Definitions of the fields in the submission manifest file name pattern

Example File Name:

OR_SUVI-L0_g16_s20171031000000_e20171031010129_v0_0_1.tar.gz.md5

Archive Ingest

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

Receipt Verification:

L0 files arrive from LZSS with a manifest file that contains a SHA-384 checksum. (These are different manifests than the MD5 ones NCEI creates.) NCEI will pull these data from LZSS using a CAC-authenticated VPN connection. NCEI will then verify files using the SHA-384 checksums. Next, NCEI will create an aggregated version of the instrument data. Tared versions of both the raw data and reformatted data will be ingested using the NCEI-Colorado Agile team's common ingest (CI) software.

Error Reconciliation:

NCEI automatically tracks problems or errors with file integrity, file name, checksum validation, or other errors that inhibit the L0 data ingest and archive. These errors are reported to LZSS, and NCEI tries to pull data a second time to complete the record. At times NCEI requests LZSS staff create new checksums if data files have missing or incorrect checksums.

Receipt Confirmation:

None.

Quality Assurance:

NCEI examines SHA-384 checksums and compare these to the checksums of the associated data files. In addition, issues with the file contents are occasionally uncovered during the aggregation and reformatting of LZSS files. NCEI also runs daily scripts to identify L0 file gaps and reports these to LZSS.

Archive File Packaging:

NCEI tars and gzips files before archival.

Archive Storage

Archive attributes of each archived file type.

Archive File Type Name: Raw L0 instrument data, tarred and gzipped; archive filename/type will be same as the original submission package.	
Archive File Attributes/IDs:	
Attribute/ID Type	Value
Storage System	archive/satellite/goesr/10/data/public/[platform]/raw/[inst]/[year]/[month]/[env]_[inst]-L0_[platform]_s[start_date]_e[end_date]_v[version].tar.gz

Archive File Type Name: Reformatted and aggregated L0 instrument data, tarred and gzipped; archive filename/type will be same as the original submission package.

Archive File Attributes/IDs:

Attribute/ID Type	Value
Storage System	archive/satellite/goesr/10/data/public/[platform]/reformat/[inst]/[year]/[month]/[env]_[inst]-L0_[platform]_s[start_date]_e[end_date]_reformat_v[version].tar.gz

Archive File Type Name: Raw L0 ISO-SERIES metadata files, tarred and gzipped; archive filename/type will be same as the original submission package.

Archive File Attributes/IDs:

Attribute/ID Type	Value
Storage System	archive/satellite/goesr/10/data/public/iso_series/[inst]/[year]/[month]/[env]_[inst]-L0-ISO-SERIES_v[version].tar.gz

Archive Updates

If errors are identified with existing AIPs, the updated copy will be archived along side the existing copies if the file with errors has been disseminated. Any such updates to files will be appropriately documented and redistributed. It is important to preserve copies of all disseminated L0 data, even files containing errors, to maintain version control and traceability in the archive.

Note that all file names include a version number. This field ensures that names of reprocessed files are always distinct from the originals.

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.

(Notional) Disposition: Unknown/TBD

Constraints

Constraint Type	Description
Use	Data acquired during Post-launch Testing (PLT) are to be considered preliminary test data.

User Community

Data will be used mostly by the NCEI-CO Space Weather Team, the Space Weather Prediction Center (SWPC) and instrument vendors. External access to the operational data should be permitted via the data manager and eventually an automated system.

User Documentation and Metadata

ISO-series metadata records will be archived as part of the AIP.

Representation Information Items

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.

Item	Description
Product Definition and User's Guide (PUG) (DOC Control # 7035538, Volume 2)	V1 describes L0 file formats
EXIS Telemetry & Command Handbook (CDRL 43, LASP doc 109743revG)	Describes EXIS packet structure (APIDs). NOTE: NOT restricted by ITAR/copyright regulations.
Flight Telemetry & Command Handbook (FTCH) For The GOES-R Space Environment In-Situ Suite (SEISS) (CDRL 43, ATC doc SEISS-D-SY043)	Describes SEISS packet structure (APIDs). NOTE: use may be restricted by ITAR/copyright regulations.
SUVI Flight Telemetry & Command Handbook (CDRL 43, Lockheed Martin doc SUVP-RQ-08-0858)	Describes SUVI packet structure (APIDs). NOTE: use may be restricted by ITAR/copyright regulations.
Magnetometer Flight Telemetry & Command Handbook (CDRL 43)	Describes MAG packet structure (APIDs). NOTE: use may be restricted by ITAR/copyright regulations.

Preservation Descriptive Information Items

Preservation Descriptive Information items contain context, provenance, and/or quality information for the data.

Item	Description
GOES-R Level Zero Storage System (LZSS) User's Guide (416-R-LZSSUSG-0283)	Provides context and information on provenance of data, along with checksum information.

Access and Dissemination

The Archive will provide access services for the data and supporting information to the designated user community.

External partners such as the Space Weather Prediction Center (SWPC) and instrument vendors such as the Laboratory for Atmospheric and Space Physics (LASP) routinely require L0 data access. NCEI is currently using a restricted FTPS server, ftps.ngdc.noaa.gov, to provide access to these organizations. It may expand this access to other authorized organizations and/or protocols in the future.

Data access for other customers will initially be provided through the data manager but may eventually move to an automated access mechanism.

Additional Terms

No additional terms.

Additional Comments:

1. Some L0 data files have been archived through the process described above. However, in the fall of 2017 NCEI turned off the GOES-R space weather common ingest stream. It was causing issues that affected not only GOES-R ingest but the ingest of other Common Ingest data streams. Considerable work has been done to make Common Ingest more robust, and it is hoped that the archiving can resume in the near future. In the meantime, data are being preserved

on spinning disk on the NFS mount nexenta1-nlsas1-vip.b:/volumes/nlsas1/goesr_private.

2. LZSS stores spacecraft orbit and attitude instrument calibration files (Data short name "SAT-INST-CAL-EPH") that are not covered by this SA. However, these same files are also included in the L1b data stream being archived in CLASS. Thus NCEI is not archiving these files with the L0 data.