



## GHRC DAAC Dataset Levels of Service Document

### Introduction

Various data collections at the GHRC DAAC may be handled with different levels of service (LoS). For some aspects of data services, such as ingest method, LoS corresponds to characteristics of the data. For other aspects of data services, LoS will depend on overall data handling priority assigned to the general categories of GHRC data holdings, specified in Table 1.

**Table 1. GHRC general data categories, with their priorities for publication**

Priority	GHRC DATA CATEGORIES
<b>SATELLITE MISSIONS</b>	
1	NASA satellite datasets (OTD, TRMM LIS, ISS LIS, AMSU)
1	Airborne validation datasets (LIP, multiple campaigns)
2	Ground validation datasets – open access (LMA)
3	Other satellite datasets (DMSP OLS, NOAA MSU)
5	Ground validation datasets – commercial, restricted access (Vaisala/NLDN, WWLLN, ENGLN)
<b>MEaSUREs PROGRAM</b>	
1	DISCOVER (RSS)
<b>FIELD CAMPAIGNS and EARTH VENTURES (Hurricane Science or GPM-GV)</b>	
1	NASA research instruments (airborne or ground, NASA-sponsored PI)
2	Affiliated research instruments (e.g., from partner university)
3	Other agency research instruments (e.g., sponsored by NOAA, DOE)
4	Ancillary research data (e.g., PERSIANN, TRMM flood maps)
5	Other agency operational data (e.g., GOES imagery, NWS radar)
<b>NASA APPLICATIONS Research Results</b>	
1	Applications products (e.g., SANDS analysis products)
3	Selected input products (e.g., MODIS subsets for selected storms)

## DAAC Data Services

### Archive

For all GHRC datasets, two local copies are maintained online, one on a public data server and another on the archive, a network attached storage device. A third, off-site backup copy is also provided on a best effort basis, as shown in Table 2. How and whether to implement this additional service is determined on a case-by-case basis, considering estimated cost of reprocessing and availability of raw data elsewhere, existence of additional copies of the data at other institutions, science value, and overall service priority among GHRC data holdings.

**Table 2. Off-site backup solutions for the different categories of GHRC data**

GHRC DATA CATEGORIES		OFF SITE BACKUP SOLUTION
<b>SATELLITE MISSIONS</b>		
1	NASA satellite datasets	Tape, manually generated (LIS and OTD); others TBD
1	Airborne validation datasets	- TBD -
2	Ground validation datasets	- TBD -
3	Other satellite datasets	- TBD -
5	Validation data - commercial	Source institution
<b>MEaSURES PROGRAM</b>		
1	DISCOVER (RSS)	PI institution
<b>FIELD CAMPAIGNS and EARTH VENTURES</b>		
1	Core NASA research instr	- TBD -
2	Affiliated research instr	- TBD -
3	Other agency research instr	- TBD -
4	Ancillary research data	- TBD -
5	Other agency operational data	Source institution
<b>NASA APPLICATIONS Research Results</b>		
1	Applications products	- TBD -
3	Selected input products	Source institution

### Ingest

The GHRC data ingest process includes acquiring data files from a provider and staging them to both public data server and archive. In some cases, additional processing may be required before the data is ready to stage (see “Post-Ingest Processing”). GHRC data ingest software routines all follow one of a few patterns, based on ingest method. An ongoing dataset will require software for automated ingest while a smaller closed dataset will only need a one-time upload. A general

rule of thumb is that any ingest process that will be repeated should be automated. For a one-time ingest, software should be implemented only when the level of effort to provide a software solution is less than the level of effort needed to ingest and stage the data manually.

### Post-Ingest Processing

Often datasets require minor processing such as renaming to better fit the GHRC data file naming convention. For better interoperability with tools and other data, translating to a standard, self-describing format like netCDF or HDF-EOS may also be needed. (Such reformatting will include a file name change if needed.) Generally reformatting is not done for data with a low LoS. Science product generation is typically done by the data provider, but in some cases we partner with them to provide science processing as well. *In this and all subsequent tables, "-" indicates that the service is not provided.*

**Table 3. Post-ingest processing for the different categories of GHRC data**

GHRC DATA CATEGORIES		POST-INGEST PROCESSING		
		Rename	Reformat	Science Proc
<b>LIGHTNING (LIS mission)</b>				
1	NASA satellite datasets	As needed	As needed	As needed
1	Airborne validation datasets		As needed	-
2	Ground validation datasets		As needed	Y
3	Other satellite datasets		As needed	As needed
5	Validation data - commercial		-	-
<b>MEaSURES PROGRAM</b>				
1	DISCOVER (RSS)		Y	-
<b>FIELD CAMPAIGNS and EARTH VENTURES</b>				
1	Core NASA research instr	As needed	As needed	As needed
2	Affiliated researcher instr	As needed	As needed	As needed
3	Other agency research instr	As needed	As needed	-
4	Ancillary research data	As needed	-	-
5	Other agency operational data	-	-	As needed*
<b>NASA APPLICATIONS Research Results</b>				
1	Applications products	As needed	-	-
3	Selected input products	-	As needed	As needed**

\* For example, say the lead PI(s) wants some non-standard product from WSR-88D data and thus works with GHRC to derive the product for all pertinent radars.

\*\* Processing may be needed to prepare input data for input to target product generation.

## Metadata and Documentation

All data held at GHRC will be cataloged with core metadata for data discovery and metrics tracking. In most cases, a data citation including author names, dataset title and digital object identifier (DOI) will be created. Of the two broad categories of documents listed here, a README is understood to provide at least the minimal information needed to identify contents of a data file. A Guide document follows a specified format and contains additional information including a brief description of the instrument and project or campaign. Additional documents such as algorithm description may be supplied by the data provider and will also be cataloged with the data.

**Table 4. Metadata and documentation required for publication of GHRC datasets, by data category**

GHRC DATA CATEGORIES		METADATA AND DOCUMENTATION			
		Catalog (discovery)	DOI and Citation	README	Guide
<b>SATELLITE MISSIONS</b>					
1	NASA satellite datasets	Y	Y	Y	Y
1	Airborne validation datasets	Y	Y	Y	Y
2	Ground validation datasets	Y	Y	Y	Y
3	Other satellite datasets	Y	Y	Y	Y
5	Val. data - commercial	Y	Y	-	-
<b>MEaSURES PROGRAM</b>					
1	DISCOVER (RSS)	Y	Y	Y	Y
<b>FIELD CAMPAIGNS and EARTH VENTURES</b>					
1	Core NASA research instr	Y	Y	Y	Y
2	Affiliated researcher instr	Y	Y	Y	Y
3	Other agency research instr	Y	As needed	Y	-
4	Ancillary research data	Y	As needed	Y	-
5	Other agency ops data	Y	As needed	-	-
<b>NASA APPLICATIONS Research Results</b>					
1	Applications products	Y	Y	Y	Y
3	Selected input products	Y	As needed	Y	-

## Distribution Services

With very few exceptions, all GHRC data are publicly available for download via FTP and HTTPS. Use of the HTTPS protocol allows GHRC to track user accesses via the ESDIS Earthdata Login Service. Depending on the dataset, its priority, and user community needs, other data access, visualization and exploration services are provided.

**Table 5. Distribution services offered for the different categories of GHRC data**

GHRC DATA CATEGORIES		DISTRIBUTION SERVICES		
		Public FTP/HTTPS	Restricted Access	Add'l Data Services
<b>SATELLITE MISSIONS</b>				
1	NASA satellite datasets	Y	-	Options include: OPeNDAP, LIS Space/time search, trends graph, vis tools
1	Airborne validation datasets	Y	-	
2	Ground validation datasets	Y	-	
3	Other satellite datasets	Y	-	
5	Validation data - commercial	-	Y*	
<b>MEaSURES PROGRAM</b>				
1	DISCOVER (RSS)	Y	-	OPeNDAP, RASI
<b>FIELD CAMPAIGNS and EARTH VENTURES</b>				
1	Core NASA research instr	Y	-	WMS or other map / vis tools as appropriate
2	Affiliated researcher instr	Y	As requested**	
3	Other agency research instr	Y	As requested**	
4	Ancillary research data	Y	As requested**	
5	Other agency operational data	Y	-	
<b>NASA APPLICATIONS Research Results</b>				
1	Applications products	Y	-	As needed
3	Selected input products	Y	-	As needed

*\* Some lightning datasets are licensed from commercial data providers by the MSFC Lightning Science team, for calibration and validation of satellite-derived lightning observations. These data are ingested and archived at GHRC in collaboration with the Lightning Science team, and access is restricted to cal/val team members.*

*\*\* Some experimental data products may be acquired for use by field campaign investigators during the campaign. These data may be archived as part of the complete field campaign collection, but not distributed beyond the science team, if so requested by the data provider.*