NSIDC DAAC Guidelines for Preliminary Metadata Creation and Data Product Delivery v2.01

Background

This document describes the preliminary metadata files and data transfer manifest we require with data product deliveries. We refer to preliminary metadata files as “premets” and “spatials”, and both need to

be generated for every file in your data product. The data transfer manifest is a single form that seeks to glean basic information about a data product to help the DAAC assure the integrity of individual files, and the delivery as a whole, following data transfer to us.

Data publication at the NSIDC DAAC is a collaborative process between you—the data producer—and us. Depending on your data product, there may be file-level attributes for which standardized keywords are needed in the premet files (see Section 1.2 below), and if so, contact your publication team lead here and we'll provide the proper keyword. Also, if you need guidance in choosing the best type of spatial file to use for your data, we can advise you (described in Section 2 below). Lastly, we can help determine the best data transfer method of your finalized data product if necessary (see Section 3 below).

When you're ready to generate premet and spatial files, and then afterward when you're ready to deliver your finalized data and ancillary files, please reach out to your publication team lead so we can coordinate advising you as necessary, answering any questions you have, and help make this process go smoothly for you.

A note on terminology: generally speaking, once files are ingested into our data archive system, they're referred to as "granules". A granule is described as the smallest aggregation of data that is independently managed by the data archive system. Granules can comprise either a single science file (e.g., one .nc, or one .tif file), or multiple science files of the same or different data types (e.g., two or more netCDF files; one or more shapefiles since they are composed of at least a .shp, .shx, and a .dbf file; or any mix of varying science file types, such as .nc, .jpg, .tif, .txt, and shapefiles). A granule containing more than one science file is referred to as a multi-file granule.

1. Premet Input Files

Premets are ASCII text files that contain file-level temporal coverage and/or standardized keywords which enable users to search for individual files within a data product based on those attributes. A premet file must be generated for each data file in a data product.

1.1 Premet File Naming Convention Requirements

The premet files must be named exactly as the data files they'll accompany, **including the data file extension*** and ending with the suffix ".premet"
**unless your data product's science files are geared to be combined to form multi-file granules when ingested. In this case the data type extension should be omitted*.

Example premet file naming convention

Single file:
A premet "ILATM2_20141023_181210_smooth_nadir3seg_50pt.**csv**.premet" takes its name straight from the single file named "ILATM2_20141023_181210_smooth_nadir3seg_50pt.csv", plus the addition of the .premet suffix.

Multi-file:
A premet named “TSX_W70-55N_19Nov13_30Nov13_09-56-46_v02_0.premet” is derived from only the shared elements of the multiple file names listed below which will be distributed together as one, multi-file granule:

```
TSX_W70-55N_19Nov13_30Nov13_09-56-46_vy_v02_0.tif
TSX_W70-55N_19Nov13_30Nov13_09-56-46_vx_v02_0.tif
TSX_W70-55N_19Nov13_30Nov13_09-56-46_vv_v02_0.tif
TSX_W70-55N_19Nov13_30Nov13_09-56-46_v02_0.txt
TSX_W70-55N_19Nov13_30Nov13_09-56-46_ey_v02_0.tif
TSX_W70-55N_19Nov13_30Nov13_09-56-46_ex_v02_0.tif
```

1.2 Premet File Content and Examples

Required premet content for any data (applicable to all data producers):

1. LocalVersionID is required and is generally set and left at a value of 001. Unlike a data set version id, this file-level LocalVersion ID isn’t a highly visible attribute for users to see and isn’t searchable with Earthdata Search. If setting it to a value other than 001 would be meaningful to you as the data producer, you may do so.
2. Temporal information is also required and should directly reflect range of dates and times covered by data within the file, or the single date and time at which the data were captured (e.g., camera images).

Required premet content only for specific data use cases (not applicable to all data producers):

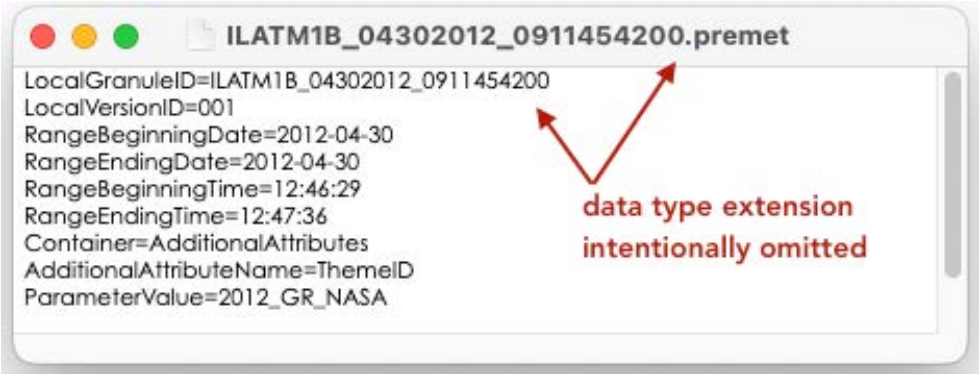
1. If files within your data product are collected during deployments (a.k.a. “campaigns” which typically relate to collection seasons and/or regions that can vary for the files within a given data product), the three attributes shown in red in Examples 2 and 3 in the table below: Container, AdditionalAttributeName, and ParameterValue, should be used.
    a. The ‘ParameterValue’ value, shown in ***bold red italic*** in both examples needs to be a standardized keyword—contact us if you need it and we'll provide you with the proper keyword.
2. If your data will have multiple science files that should be distributed as a multi-file granule, a LocalGranuleID attribute is necessary in the premet. The value of this attribute shown in Example 4 in **bold purple** is the shared portion of the multiple science file names—the science file extensions **must not be included**. In the case of the following two science files: BLATM2_20021203_124629_icessn_summary_50pt, and BLATM2_20021203_124629_smooth_nadir3seg_50pt, "BLATM2_20021203_124629" would end up being the multi-file granule name, and is what’d be displayed in Earthdata Search's granule listing. A multi-file granule serves as a container for the two science files—the files it contains are not renamed. When users download a multi-file granule, they receive both science files.
3. If you'd like a browse image to be displayed with each science file in Earthdata Search’s granule view, the LocalGranuleID attribute also needs to be included in the premet file, but in this case should comprise the full science file name, including extension, as shown in **bold green** in Example 5 in the table below.

Example 1: Premet content for a single-file granule covering a date and/or time range	Example 2: Premet granule for a single-file granule where ThemeID is relevant
LocalVersionID=001	LocalVersionID=001
RangeBeginningDate=2011-09-25	RangeBeginningDate=2012-03-15
RangeEndingDate=2011-10-01	RangeEndingDate=2012-03-15
RangeBeginningTime=00:23:10	RangeBeginningTime=18:52:34.00
RangeEndingTime=23:59:56	RangeEndingTime=23:17:25.10
	Container=AdditionalAttributes
	AdditionalAttributeName=ThemeID
	*ParameterValue=**2012_GR_NASA***
Example 3: Premet granule for a single-file granule wherein data are recorded at a single point in time and ThemeID is relevant	**Example 4: Premet point for a multi-file granule**
LocalVersionID=001	LocalVersionID=001
RangeBeginningDate=2009-10-25	**LocalGranuleID=BLATM2_20021203_124629**
RangeBeginningTime=16:27:01	RangeBeginningDate=2002-12-03
Container=AdditionalAttributes	RangeEndingDate=2002-12-03
AdditionalAttributeName=ThemeID	RangeBeginningTime=12:46:29
ParameterValue=***2009_AN_NASA***	RangeEndingTime=12:47:36
Example 5: Premet point for a single-file granule with a browse image	
LocalVersionID=001	
LocalGranuleID=IS2ATBABD_20190501_20210930_v01.gpkg	
RangeBeginningDate=2019-05-01	
RangeEndingDate=2021-09-30	
RangeBeginningTime=00:00:00.000	
RangeEndingTime=23:59:59.999	

Example premet files

1. The following image shows the attributes included in a premet for a data product comprising single-granule files (i.e., one granule = one file) for which ThemeID isn't relevant.

2. The image below shows an example of a premet for a data product comprising multi-file granules (i.e., one granule = multiple files) for which ThemeID *is* relevant:



Things to note for multi-file granule premets: 1) a LocalGranuleID attribute is required to define the file name, however the file extension *should be omitted* both from that attribute AND the name of the premet file.

2. Spatial Input Files

Spatial files are ASCII text files that define file-level spatial extents. This information enables users to search for individual files in a data product based on spatial coverage. A spatial file must be generated for each data file in your data product.

There are a number of options for representing the spatial extents of data files, those being: points, bounding rectangles or irregular polygons.

- Data within a file collected at a discrete lat/lon location is typically represented as a point.
- Rectangular bounding extents can be useful for any files containing data that fit neatly within rectangular extents (e.g., gridded data).
- Irregular polygons are often used to define meandering data extents (e.g., airborne data collected along flight lines).

2.1 Spatial File Naming Convention Requirements

Just as premets do, spatial files need to be named exactly as the data files they'll accompany (***including the data file extension***, except in the case of multi-file granules as noted above in the Premet File Naming Conventions), and be suffixed with either ".spatial" or ".spo".

The .spatial file extension is used for data files containing:

- data collected at a single (point) location
- content (like gridded/raster) that is best represented by a bounding rectangle
- vertices of data collected along an irregular path;

The ".spo" file extension is used for data files where you are providing vertices defining a polygon boundary around a given file's data point cloud content.

Example .spatial / .spo naming convention:

Single-file granule examples:
ILATM2_20120328_142406_smooth_nadir3seg_50pt.**csv**.spatial
DMS_1842639_09308_20180407_23033937.**tif**.spo

Multi-file granule example:
ILATM1B_04302012_0911454200.spatial
TSX_W70-55N_19Nov13_30Nov13_09-56-46_v02_0.spo

2.2 Example Spatial Files

2.2.1 Point Data Spatial File Requirements

The file must contain one coordinate pair. Longitude is required to be listed first, and should be specified from -180 to 180 decimal degrees, followed by whitespace, and then latitude specified from -90 to 90 decimal degrees. The coordinates can have up to 6 decimal places.

Example point data spatial file contents

The following image shows a spatial file specifying a point at -49.565488 degrees longitude, and 69.184703 degrees latitude.

2.2.2 Bounding Rectangle Spatial File Requirements

The file must contain two coordinate pairs. The coordinate pair in the first line describes the northwest corner of the bounding rectangle, longitude followed by latitude. The coordinate pair on the second line describes the southeast corner of the bounding rectangle, longitude followed by latitude. The coordinates can have up to 6 decimal places.

Example bounding rectangle spatial file - contents and plot #1

The following image shows a spatial file specifying a bounding rectangle with the northwest corner at -10.00 degrees longitude, 10.00 degrees latitude, and the southeast corner at -5.00 degrees longitude, 5.00 degrees latitude.

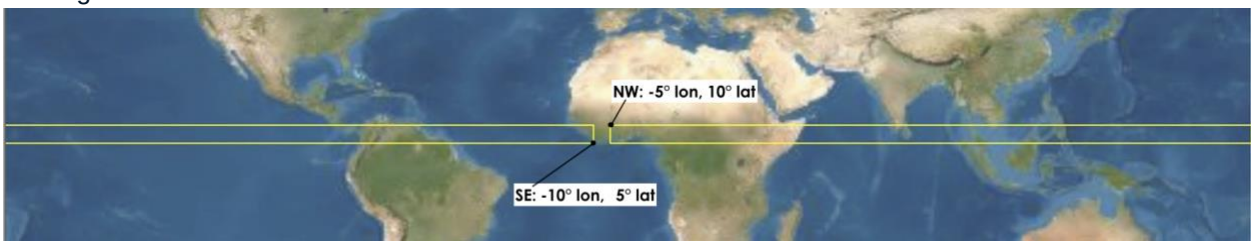
Plot 1 displays the bounding rectangle coordinates in black, and in yellow how the data file's post-publication boundary would look:

Plot 1

Example bounding rectangle spatial file - contents and plot #2

The following image shows a spatial file specifying a bounding rectangle with the northwest corner at -5.00 degrees longitude, 10.00 degrees latitude, and the southeast corner at -10.00 degrees longitude, 5.00 degrees latitude. This leads to a rectangular band which nearly encircles the Earth.

Plot 2 displays the bounding rectangle coordinates in black, and in yellow how the data file's post-publication boundary would look - you can see the breadth of the area encompassed by this bounding rectangle:



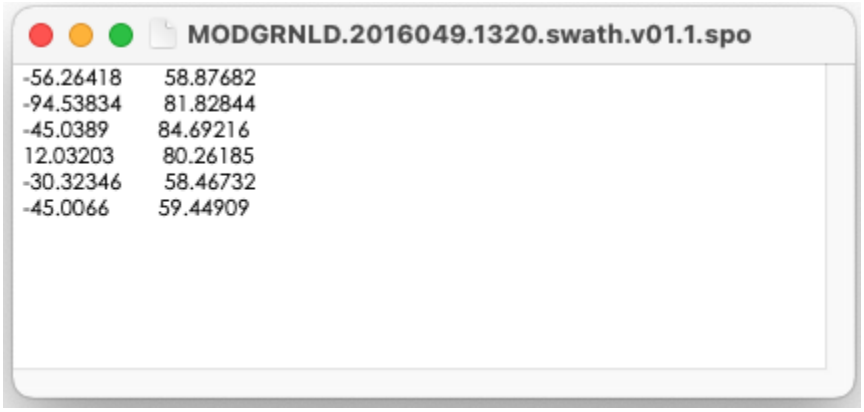
Plot 2

2.2.3 Polygon/Irregular Polygon Spatial File Requirements

Longitude is required to be listed first, and should be specified from -180 to 180 decimal degrees, followed by whitespace, and then latitude specified from -90 to 90 decimal degrees. The coordinates can have up to 6 decimal places.

Example irregular polygon .spo file - content and plot 3

The following image shows a .spo-type spatial file specifying vertices that define a polygon. The vertices defined by the coordinate pairs must be listed in a clockwise direction as viewed from above the Earth's surface:

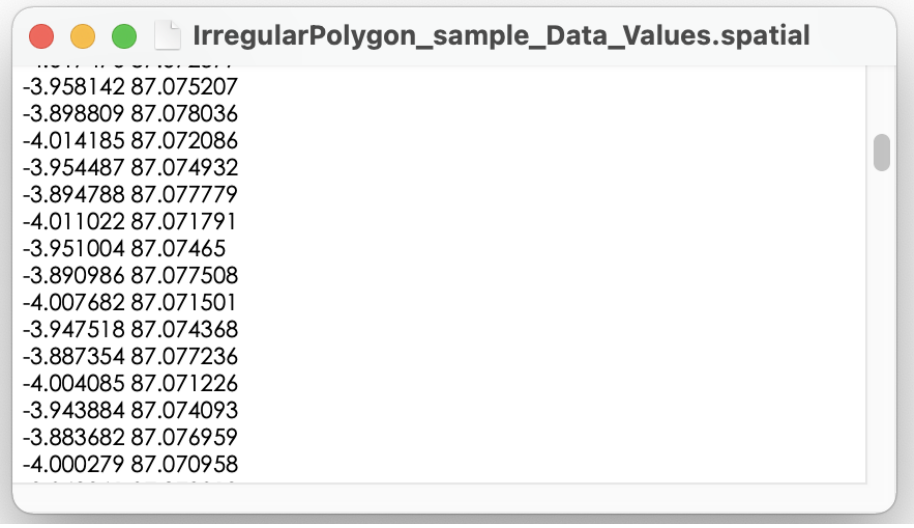


Plot 3 displays the polygon coordinates in pink, and outline in green how the data file's post-publication boundary would look:

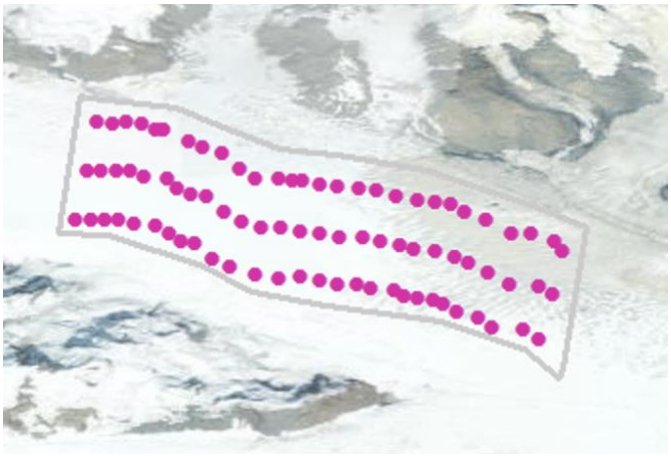
Plot 3

Example irregular polygon .spatial file - content and plot 4

This example shows a .spatial-type spatial file that lists data point cloud coordinate values collected along an airborne survey segment by, in this case, a conically-scanning lidar system.



Plot 4, below, displays the data coordinates in pink, and in gray approximately how the data file's post-publication boundary would look:



Plot 4

3. Delivering Data to NSIDC

Once sample data review is complete and file naming, format, and structure of your data have been finalized along with the creation of premets and spatials, we offer a number of options for you to transfer your data product to us - shipping data via hard drives is not one of those options. When you are ready, please complete a Data Transfer Manifest and reach out to your NSIDC DAAC mission/project lead to work with you to determine the most suitable data transfer method for your data.

3.1 Data Transfer Manifest

A completed data transfer manifest (see example below) must accompany every batch of data transferred to the DAAC to facilitate the ingest of your data in a timely manner. It allows us to verify that all files expected to be delivered have been received and were not corrupted during transfer.

Example completed data transfer manifest

NSIDC DAAC Data Transfer Manifest

Operations contact email: ops@nsidc.org
Instructions: Please fill out the following and include this with the data product being transferred.
Thank you for your help in completing this thoroughly!

Data Producer Name: <name of best data delivery point of contact>
Organization Name: Lamont-Dohert Earth Observatory
Contact Email Address: bestcontact@email.here
Date of transfer: 4 may 2015
Data Product Title: IceBridge Sander AIRGrav L1B Geolocated Free Air Gravity Anomalies
Data Product Short Name (if known; e.g., NSIDC-0451, SNEX21_DSM, etc.): IGGRV1B
Campaign (Region/year, if applicable): Antarctica 2012

Is this the first delivery of this Data Set? ☐ Yes ☒ No

Has the data format changed since the most recent delivery?
☐ Yes ☒ No If Yes was selected, please describe:

Has the file naming convention changed since the most recent delivery?
☐ Yes ☒ No If Yes was selected, please describe:

If users will require an updated reader to view these files, will it be included in this delivery?
☐ Yes ☐ No ☒ N/A

Data Type(s), (e.g., nc, csv, h5, etc.): ASCII, premet, spatial
Total number of files expected for each data type: 42 of each

Are browse files included?
☐ Yes ☒ No
If yes, browse file types must be either jpg, png, or tif
(however Earthdata Search does not render tif browse images).
Browse file type: ☐ jpg. ☐ .png ☐ .tif Count:

Total volume of data being transferred (include units):
Type of checksum (MD5 or CKSUM): CKSUM

Please include here a comma-separated list of the following for each file that is being delivered:
FileName, FileSize, File Checksum
IGGRV1B_V01_20121012.xyz, 25109630, 1246638426
IGGRV1B_V01_20121012xyz.premet, 470, 3768648749
IGGRV1B_V01_20121012.xyz.spatial, 128000, 2642184452
IGGRV1B_V01_20121013.xyz, 25999330, 4259220357
IGGRV1B_V01_20121013.xyz.premet, 475, 958200890 … (etc.)

Change Log

Major version changes (advancing from 1.x to 2.x) are those that significantly impact the intended audience.

Minor version changes (advancing from 1.x to 1.x+1) are those that do not significantly impact the intended audience and, in general, enhance or clarify details within the document (e.g., update a reference to another document, update a point-of-contact).

Revision	Effective Date	Description of Changes	Author(s)
2.01	22 Feb 2024	Within the Premet File Content Examples section, updated to specify using LocalGranuleID for single-file granules with browse images to hardcode the original file name to show as the granule's name in EDSC. Fixed image sizes after cloud conversion removed original size settings.Updated content of premet example images.	Amy FitzGerrell
2.0	28 Nov 2023	Revisions made to discontinue requirement that platform, instrument, sensor, and aircraft/tail ID attributes from premet input files.	Amy FitzGerrell
1.0	26 Oct 2022	Original document	Amy FitzGerrell