

NASA SDS Product Specification

Level-1 Range Doppler Single Look Complex

L1 RSLC

Rev D

JPL D-102268

May 30, 2024, Version 1.1.2

Author: Brian Hawkins

Copies of this document obtained outside of JPL's Product Lifecycle Management (PLM) tool may not be current. Please contact Heresh Fattahi [\(heresh.fattahi@jpl.nasa.gov\)](mailto:heresh.fattahi@jpl.nasa.gov) to determine if a more recent revision is available.

National Aeronautics and Space Administration

Jet Propulsion Laboratory California Institute of Technology Pasadena, California

SIGNATURE PAGE

Prepared by:

Electronic Signature on File 07-Jul-2024 Brian Hawkins, NISAR ADT Level-1 Product Lead Date

Approved by:

Electronic Signature on File 02-Jul-2024 Ana Maria Guerrero, NISAR Mission System Manager Date

Electronic Signature on File 08-Jul-2024 Chuck Baker, NISAR Mission System Engineer Date

Electronic Signature on File 09-Jul-2024 Paul Rosen, NISAR Project Scientist Date Date

Electronic Signature on File 02-Jul-2024 Cecilia Cheng, NISAR SDS Manager Date

Electronic Signature on File 02-Jul-2024 Heresh Fattahi, NISAR ADT Lead Date

Electronic Signature on File 09-Jul-2024 Helen Mortensen, NISAR SDS Lead System Engineer Date

EPDM ELECTRONIC SIGNATURES

ACKNOWLEDGEMENT

The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).

© 2024 California Institute of Technology. Government sponsorship acknowledged.

DOCUMENT CHANGE LOG

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF FIGURES

1 INTRODUCTION

1.1 Purpose of Description

This document provides a specification of the NASA-ISRO Synthetic Aperture Radar (NISAR) L-SAR Level-1 Range Doppler Single Look Complex (RSLC) product to be generated by the NASA Science Data System (SDS) and provided to the Distributed Active Archive Center (DAAC). This data product is usually referenced by the short name RSLC.

1.2 Document Organization

Section 2 provides an overview of the product, including its purpose, and latency.

Section 3 provides the structure of the product, including granule definition, file organization, spatial resolution, temporal and spatial organization of the content, the size and data volume.

Section 4 provides qualitative descriptions of the information provided in the product.

Section 5 provides a detailed identification of the individual fields within the RSLC product, including for example their units, size, and coordinates.

Section 6 provides a description of the metadata cube representation.

Appendix A provides a listing of the acronyms used in this document.

1.3 Applicable and Reference Documents

Applicable documents levy requirements on areas addressed in this document. Reference documents are cited to provide additional information to readers. In case of conflict between the applicable documents and this document, the Project shall review the conflict to find the most effective resolution.

Applicable Documents

[AD7] ISO-19115-2,<https://www.iso.org/obp/ui/#iso:std:iso:19115:-2:ed-2:v1:en>

Reference Documents

- [RD1] NISAR NASA SDS Algorithm Theoretical Basis Document, JPL D-95677, Rev A, November 12, 2023.
- [RD2] EOSDIS Handbook, July 2016, retrieved from [https://cdn.earthdata.nasa.gov/conduit/upload/5980/EOSDISHandbookWebFinaL2.pdf](https://cdn.earthdata.nasa.gov/conduit/upload/5980/EOSDISHandbookWebFinal1.pdf)
- [RD3] NISAR SDS File Naming Conventions, JPL D-102255, Rev A, April 28, 2023.
- [RD4] HDF5 documentation at<https://portal.hdfgroup.org/display/HDF5/HDF5>
- [RD5] Eineder, M. (2003), Efficient simulation of SAR interferograms of large areas and of rugged terrain, IEEE Transactions on Geoscience and Remote Sensing, 41(6), 1415-1427.
- [RD6] NASA SDS Radar Pointing Product Software Interface Specification, JPL D-102264, Rev B, November 12, 2020

2 PRODUCT OVERVIEW

2.1 Product Background

Each NASA SDS L0B-L2 LSAR product [\(Figure 2-1](#page-10-2) and [Table 2-1](#page-11-0) Product Dependency) is distributed as a single Hierarchical Data Format version 5 (HDF5) [\[RD4\]](#page-9-0) granule. All the metadata and imagery data are packaged in clearly defined sub-groups within the granule in compliance with the HDF5 specification. The NISAR product level definitions are given in [Table 2-2.](#page-12-1)

Figure 2-1 Product Dependency

Table 2-1. Key to Product Dependency Diagram

2.2 RSLC Overview

The RSLC product is in the zero-Doppler radar geometry convention [\[RD1\].](#page-9-1) The output image is on a grid characterized by constant azimuth time interval and one-way slant range spacing. The output grid is also characterized by a fixed set of starting slant range, azimuth time interval, and slant range spacing values to allow for easy interpolation. All the primary image layers for a multi-polarization or multi-frequency product are generated on a common azimuth time-slant range grid.

The RSLC product, which is used to derive other L1/L2 products, contains individual binary raster layers representing complex signal return for each polarization layer. The RSLC data corresponding to the auxiliary sub-band is stored in a similar format but in a separate data group within the HDF5 product granule. The RSLC product is also packed with input, instrument and processing facility information; processing, calibration and noise parameters; geolocation grid; and data quality flags.

The RSLC product complex backscatter is in Digital Numbers (DNs) with secondary layer look up tables (LUTs) provided to convert to beta-naught, sigma-naught, and gamma-naught.

These radiometric correction LUTs are defined with respect to the ellipsoid (e.g, not with respect to the local terrain). Additional secondary layers of slowly varying quantities are compactly stored in metadata cubes (see Sec [6\)](#page-53-0).

All standard (i.e., non-urgent response) products are processed using the Medium-fidelity Orbit Ephemeris (MOE) product for forward processing and the Precise Orbit Ephemeris (POE) product for reprocessing campaigns.

The RSLC product groups with their basic properties are given in Section [4.](#page-24-0) The details of the data elements are given in Section [5.](#page-29-0) Metadata cubes are discussed in Section [6.](#page-53-0)

3 PRODUCT ORGANIZATION

3.1 File Format

All NISAR standard products are in the Hierarchical Data Format version 5 (HDF5) [\[RD4\].](#page-9-0) HDF5 is a general-purpose file format and programming library for storing scientific data. The National Center for Supercomputing Applications (NCSA) at the University of Illinois developed HDF to help scientists share data more easily. Use of the HDF library enables users to read HDF files regardless of the underlying computing environments. HDF files are equally accessible in Fortran, C/C++, and other high-level computation packages such as IDL or MATLAB.

The HDF Group, a spin-off organization of the NCSA, is responsible for development and maintenance of HDF. Users should reference The HDF Group website at <https://portal.hdfgroup.org/display/HDF5/HDF5> [\[RD4\]](#page-9-0) to download HDF software and documentation.

HDF5 represents a significant departure from the conventions of previous versions of HDF. The changes that appear in HDF5 provide flexibility to overcome many of the limitations of previous releases. The basic building blocks have been largely redefined, and are more powerful but less numerous. The key concepts of the HDF5 Abstract Data Model are Files, Groups, Datasets, Datatypes, Attributes and Property Lists. The following sections provide a brief description of each of these key HDF5 concepts.

3.1.1 HDF5 File

A File is the abstract representation of a physical data file. Files are containers for HDF5 Objects. These Objects include Groups, Datasets, and Datatypes.

3.1.2 HDF5 Group

Groups provide a means to organize the HDF5 Objects in HDF5 Files. Groups are containers for other Objects, including Datasets, named Datatypes and other Groups. In that sense, groups are analogous to directories that are used to categorize and classify files in standard operating systems.

The notation for files is identical to the notation used for Unix directories. The root Group is "/". A Group contained in root might be called "/myGroup." Like Unix directories, Objects appear in Groups through "links". Thus, the same Object can simultaneously be in multiple Groups.

3.1.3 HDF5 Dataset

The Dataset is the HDF5 component that stores user data. Each Dataset associates with a Dataspace that describes the data dimensions, as well as a Datatype that describes the basic unit of storage element. A Dataset can also have Attributes.

3.1.4 HDF5 Datatype

A Datatype describes a unit of data storage for Datasets and Attributes. Datatypes are subdivided into Atomic and Composite Types.

Atomic Datatypes are analogous to simple basic types in most programming languages. HDF5 Atomic Datatypes include Time, Bitfield, String, Reference, Opaque, Integer, and Float. Each atomic type has a specific set of properties. Examples of the properties associated with Atomic Datatypes are:

- Integers are assigned size, precision, offset, pad byte order, and are designated as signed or unsigned.
- Strings can be fixed or variable length, and may or may not be null-terminated.
- References are constructs within HDF5 Files that point to other HDF5 Objects in the same file.

HDF5 provides a large set of predefined Atomic Datatypes. [Table 3-1](#page-15-2) lists the Atomic Datatypes that are used in NISAR data products.

HDF5 Atomic		
Datatypes	Description	
H5T_STD_U8LE	unsigned, 8-bit, little-endian integer	
H5T STD U16LE	unsigned, 16-bit, little-endian integer	
H5T_STD_U32LE	unsigned, 32-bit, little-endian integer	
H5T STD U64LE	unsigned, 64-bit, little-endian integer	
H5T STD I8LE	signed, 8-bit, little-endian integer	
H5T STD I16LE	signed, 16-bit, little-endian integer	
H5T STD I32LE	signed, 32-bit, little-endian integer	
H5T STD I64LE	signed, 64-bit, little-endian integer	
H5T IEEE F32LE	32-bit, little-endian, IEEE floating point	
H5T IEEE F64LE	64-bit, little-endian, IEEE floating point	
H5T C S1	character string made up of one or more bytes	

Table 3-1. HDF5 Atomic Datatypes

Derived Datatypes are user-defined variants of predefined Atomic Datatypes where the data organization has been modified at the bit-level. Derived data types are particularly useful for representing custom N-bit integers and floating point numbers.

Composite Datatypes incorporate sets of Atomic datatypes. Composite Datatypes include Array, Enumeration, Variable Length and Compound.

• The Array Datatype defines a multi-dimensional array that can be accessed atomically.

- Variable Length presents a 1-D array element of variable length. Variable Length Datatypes are useful as building blocks of ragged arrays.
- Compound Datatypes are composed of named fields, each of which may be dissimilar Datatypes. Compound Datatypes are conceptually equivalent to structures in the C programming language.

Named Datatypes are explicitly stored as Objects within an HDF5 File. Named Datatypes provide a means to share Datatypes among Objects. Datatypes that are not explicitly stored as Named Datatypes are stored implicitly. They are stored separately for each Dataset or Attribute they describe.

NISAR products employ the following Derived and Compound Datatypes.

Table 3-2 NISAR HDF5 Derived and Compound Datatypes

3.1.5 HDF5 Attribute

An Attribute is a small aggregate of data that describes Groups or Datasets. Like Datasets, Attributes are also associated with a particular Dataspace and Datatype. Attributes cannot be subsetted or extended. Attributes themselves cannot have Attributes.

3.2 NISAR File Organization

3.2.1 Groups

All NISAR HDF5 files are organized as groups with no actual data at the root("/") level. Table 3-3 shows the general layout of the HDF5 files that are generated by the NISAR Science Data System. All data are organized under "/science" with data from the L-SAR and S-SAR instruments separated into their own groups.

Table 3-3 Group organization at the top level of a NISAR HDF5 File

In the nominal baseline, L-SAR and S-SAR data will not appear in the same granule, even if they cover the same geographic area. Data structure described below the primary groups ("/science/LSAR" for L-SAR and "/science/SSAR" for S-SAR) will be the same for L-SAR and S-SAR products. The rest of the document from this point on describes the layout of the product containing L-SAR data. The specification for equivalent S-SAR data products is expected to be the same except for the substitution of "LSAR" by "SSAR" in the dataset paths in the HDF5 granule.

3.2.2 File Level Metadata

Global metadata at the file level are currently given as Global Attributes shown in Table 3-4.

Metadata regarding the data in the particular granule are given in "/science/[L|S]SAR/identification" for L- or S-SAR. These data are described further in Sec [4.2](#page-24-2) and Sec [5.2.](#page-31-0)

3.2.3 Variable Metadata (HDF5 Attributes)

NISAR standards incorporate additional metadata that describe each HDF5 Dataset within the HDF5 file. Each of these metadata elements appear in an HDF5 Attribute that is directly associated with the HDF5 Dataset. Wherever possible, these HDF5 Attributes employ names that conform to the Climate and Forecast (CF) conventions.

Table 3-5 lists the CF names for the HDF5 Attributes that NISAR products typically employ.

Some HDF5 datasets are populated with statistical attributes. [Table 3-6](#page-18-2) and [Table 3-7](#page-18-3) describe statistical attributes added to real- and complex-valued HDF5 datasets, respectively. The list of real- and complex-valued HDF5 datasets for the standard RSLC product is given in [Table 3-8.](#page-19-4)

3.3 Granule Definition

NISAR RSLC granules will conform to the Tiling Scheme being developed for the mission and are expected to have a ground footprint of 240 km x 240 km.

3.4 File Naming Convention

NISAR RSLC Granule names will conform to the Standard Product File Naming Scheme [\[RD3\].](#page-9-2)

3.5 Temporal Organization

The RSLC data are arranged on a uniformly spaced, increasing zero-Doppler azimuth time grid. Using row-major order convention of representing 2D raster arrays, zero-Doppler azimuth time is represented by the row direction or the slowest changing dimension.

3.6 Spatial Organization

The RSLC data are arranged on a uniformly spaced, increasing zero-Doppler azimuth time in the row direction and increasing slant range grid in the column direction following the row-major order convention of representing 2D raster arrays.

3.7 Spatial Sampling and Resolution

The NISAR L-SAR uses a non-uniformly spaced sequence of pulses in SweepSAR mode to collect radar data, to overcome the limitations imposed by transmit gaps affecting the wide imaging swath [\[RD1\].](#page-9-1) Processing software accounts for the non-uniform sampling to generate the final RSLC product on a uniform grid. Some salient features of the output grid for the RSLC product are:

- 1. The center of the top-left pixel will correspond to the same zero-Doppler azimuth time and slant range for all imagery layers in an L-SAR RSLC product – frequency A and frequency B.
- 2. All imagery layers in an L-SAR RSLC product frequency A and frequency B, are generated on the same zero-Doppler azimuth time grid corresponding to a 1520 Hz PRF, which is approximately 1.2 times the processed azimuth bandwidth and results in roughly 5 m ground postings.
- 3. The slant range sampling is generally 1.2 times the range bandwidth. For example, 20 MHz data are sampled at 24 MHz. The only exceptions are 77 MHz data, which are sampled at 96 MHz.
- 4. The main (frequency A) and auxiliary (frequency B) bands of L-SAR data have an exact integer scaling relationship. All bands are sampled at an integer multiple of 6 MHz.

The RSLC products are all processed to 6 m azimuth resolution. No windowing or whitening is applied in azimuth, so the antenna pattern determines the shape of the azimuth spectrum. A Kaiser window with shape parameter 1.6 is applied in range. A nominal impulse response is shown in [Figure 3-1.](#page-20-1)

Figure 3-1 Impulse response and spectrum of simulated NISAR data (20 MHz range bandwidth and 1910 Hz dithered PRF).

3.7.1 Along Track Mosaicking

The spatial sampling of the output grid has also been designed to facilitate along-track mosaicking of contiguous RSLC product granules if the user desires. The following features simplify the implementation of along-track mosaicking

- 1. The slow time sampling frequency (inverse of the zero Doppler time spacing between consecutive lines) will be chosen to be an integer, to allow synchronization between adjacent granules at integer second boundaries without the need for resampling in the azimuth time direction.
- 2. The slant range to the first pixel will be a multiple of the lowest sampling frequency (corresponding to 5MHz) to enable concatenation of adjacent granules with simple integer shifts of imagery in the slant range direction.

3.7.2 Partially Compressed RSLC Data

Some applications can benefit from using partially compressed data in near and far ranges, as well as in transmit gaps during operation in constant Pulse Repetition Frequency (PRF) mode (see [Figure 3-2\)](#page-22-0). The number of contiguous image swaths is given by a variable named "numberOfSubSwaths". The slant range extent for each of these contiguous, fully focused regions is captured in an array named "validSamplesSubSwathN" where "N" is the index of the contiguous regions in [1,5]. Each of these extent arrays are as long as the raster imagery themselves and each line contains two numbers indicating the starting index and last index in pixels (using Python convention).

Partially compressed (processed) data should be explicitly discarded for radiometric studies and for generation of polarimetric products.

Figure 3-2 Representation of valid and partially compressed samples in constant PRF and dithered PRF modes

3.8 Cloud Optimizations

NISAR science data products utilize several special features of the HDF5 format to optimize file sizes and enable high-performance read access in a cloud environment. A key challenge of cloud data access is the latency associated with calls to the cloud storage API, so the following strategies are used to minimize the number of cloud API calls needed per byte of data read:

- Chunks: Large datasets within the products use [chunked storage.](https://portal.hdfgroup.org/documentation/hdf5-docs/advanced_topics/chunking_in_hdf5.html) Every read operation thus fetches at least one entire chunk of data. The chunk size is nominally 512x512 pixels, though the precise chunk dimensions should be obtained using the [H5Pget_chunk](https://portal.hdfgroup.org/hdf5/develop/group___d_c_p_l.html#ga4ef814034f601f48ab1ed6db79b4354c) method of the HDF5 C API (or its equivalent in other language bindings).
- Compression: Data are written using a compression filter, minimizing the amount of data stored and hence transferred over the network. The HDF5 API handles decompression automatically.
- Paging: Files are created with the "paged" file space strategy [\(H5F_FSPACE_STRATEGY_PAGE](https://portal.hdfgroup.org/hdf5/develop/group___f_c_p_l.html#ga167ff65f392ca3b7f1933b1cee1b9f70) in the HDF5 C API). These pages serve as the basic unit of allocation within the file. The page size is chosen larger than the chunk size so that both a chunk of data and its HDF5-internal metadata can be read in a single cloud API call. This parameter may be queried using the [H5Pget_file_space_page_size](https://portal.hdfgroup.org/hdf5/develop/group___f_c_p_l.html#gaab5e8c08e4f588e0af1d937fcebfc885) method of the HDF5 C API.

Software that reads NISAR products stored on the cloud should take heed of the following recommendations:

- Set the page buffer size to a multiple of the file space page size using [H5Pset_page_buffer_size](https://portal.hdfgroup.org/hdf5/develop/group___f_a_p_l.html#ga8008cddafa81bd1ddada23f6d9a161ca) in the HDF5 C API. This enables caching logic that reduces the number of cloud API calls in the file driver.
- Implement chunk-aligned data access patterns. Reads in multiples of the chunk size (and aligned with chunk boundaries) are most efficient.
- If other access patterns are desired, try setting the read cache large enough to hold all the chunks that may be re-read. For example, line-by-line access can still be efficient if the read cache is large enough to hold N lines, where N is the chunk dimension. That way lines can be read from the cache instead of fetching the same set of chunks N times over the network. The cache size may be set globally using the H5Pset cache or locally with the H5Pset chunk cache methods of the HDF5 C API.

Note that, in general, these optimizations require knowledge of the file contents. Therefore, the most robust approach is to open the file, inspect the contents (e.g., chunk size, page size, and dataset dimensions) and then re-open the file with optimal parameters.

4 LEVEL 1 SINGLE LOOK COMPLEX PRODUCT

In this section, we briefly describe the layout of RSLC data and associated metadata in the NISAR HDF5 file. Detailed description of Group and Dataset names can be found in Section 5. In this section, we focus on the organization of L-SAR instrument data under the Group name "/science/LSAR".

4.1 Shapes and Dimensions of Data

Information on the shapes and dimensions of the data items in various data tables are described as part of the metadata (Sec [5.1\)](#page-29-1). This information is useful both as part of the product identification and for setting up further processing, i.e., dimensioning arrays.

4.2 Product Identification

Information needed to identify this particular product is given under the Group "/science/LSAR/identification" (Sec [5.2\)](#page-31-0). This includes information such as orbit number, track-frame number, acquisition times, a polygon representing the bounding box of the included imagery in geographic coordinates, and product version.

4.2.1 Composite Release Identifier

The Composite Release Identifier (CRID) is a global version identifier documenting the algorithms and the overall status of the science data system used to generate the product. The CRID follows the format *EPMMmp* where:

- **E** (**Environment**): a single character representing the environment or the venue where the product was generated. It can assume the values:
	- o *A*: if the product was generated in the Algorithm Development environment
	- \circ *D*: if the product was generated in the Development environment
	- o *P*: if the product was generated in the Production environment
	- \circ *T*: if the product was generated in the Integration and Test (I&T) environment
- **P (Mission Phase)**: a single numerical digit indicating the mission phase in which the product was generated. It can assume the following values:
	- o *0*: for pre-launch (Phase D)
	- o *1*: for primary science phase operations (Phase E)
	- o *2*: extended mission (Phase E)
	- o *3*: post-operations (Phase F), decommissioning, end of mission processing
- **MM (Major Release)**: two numeric digits monotonically increasing between 0 and 99. The Major Release resets to zero upon a change in the Mission Phase identifier. A change in the Major Release indicates a major change in the products i.e., a change to one or

more algorithms or to the processing rules having a significant impact on the science content of the product. The Major Release stands as a composite of the versions of all the algorithms used in the science data production systems. Individual algorithm versions are allocated in the product metadata.

- **m (Minor Release):** a single numeric digit increasing monotonically between 0 and 1 indicating a minor update to the product and/or the data system. A change in the Minor Release identifier indicates minor algorithm changes (e.g., bug fixes, small functional updates) that do not have a significant impact on the product. The Minor Release identifier resets to zero upon every update to the Major Release identifier
- **P (Patch Release)**: a single numerical digit monotonically increasing between 0 and 1. A change in the Patch Release identifier indicates an update to the science data system software that has undergone the System Deployment Review to fix a critical bug. The Patch Release resets to zero upon updates to the Major Release or Minor Release identifiers.

4.3 Radar Imagery

All the imagery layers corresponding to the RSLC product are organized by center frequency under the Group "/science/LSAR/RSLC/swaths". For L-SAR imaging modes with split imaging bands, the data is further organized into individual groups labeled "frequencyA" and "frequencyB". Imagery layers are further organized as individual 2D datasets by polarization (TxRx) within the frequency sub-groups, i.e., dataset

"/science/LSAR/RSLC/swaths/frequencyA/HH" corresponds to the SLC imagery layer for polarization combination HH processed with center frequency corresponding to frequencyA.

The details of the data elements are given in Section [5.3.](#page-34-0)

4.4 Radar Metadata

Radar metadata needed to interpret the amplitude and phase information, as well as the geolocation of the imagery are organized under the Group "/science/LSAR/RSLC/metadata".

4.4.1 Calibration Information

The subgroup "calibrationInformation" contains two major types of information as shown in Section [5.4.](#page-38-0)

4.4.1.1 Radiometric Calibration

Secondary lookup tables (LUT), common to all frequencies and polarizations as these are purely a function of imaging geometry, are organized under the subgroup "calibrationInformation/geometry". The radar imagery themselves are provided as Digital Numbers (DNs), and LUTs are provided to transform the DNs to beta0, sigma0, and gamma0 (with respect to the reference ellipsoid) according to the following

```
beta0 = abs(RSLC)^2 / beta0 LUT^2
sigma0 = abs(RSLC)^2 / signal0 LUT^2gamma = abs(RSLC)^2 / gamma = 0 LUT<sup>2</sup>
```
These LUTs are provided as a sparse grid in radar coordinates, and values at any location can be obtained using simple 2D interpolation (bilinear or higher order). After the above LUTs are applied, the resulting values have units of m^2/m^2 corresponding to radar cross section (m^2) normalized by a reference area.

4.4.1.2 Radar Information

Complex two-way antenna patterns and noise-equivalent sigma0 (nes0) are provided organized by frequency and polarization. Noise-equivalent-sigma0 could be used to apply noise correction during radiometric calibration. These datasets are provided on a sparse grid in map coordinates and values of interest at any geographical location can be estimated using simple 2D interpolation (bilinear or higher order).

4.4.2 Processing Information

Metadata giving processing parameters, algorithms, and inputs used are given in Section [5.5.](#page-47-0)

4.4.2.1 Parameters

Common parameters such as reference terrain height and chirp weighting parameters are included in the group "processingInformation/parameters". All processing parameters that vary spatially are organized on low resolution grids, to allow for easy lookup based on radar coordinates.

4.4.2.2 Algorithm Information

The processing algorithm information is provided in the subgroup "processingInformation/algorithms/". It includes the software version ("softwareVersion"), which is the version of the ISCE3 software that was used to generate the product, and the list of algorithms employed in the product processing.

4.4.2.3 Inputs

The key input files – L0B granules, orbit, attitude, calibration, DEM source description, and configuration files are tracked and listed under the subgroup "processingInformation/inputs".

4.4.3 Other Radar Metadata

Section 5.6 includes the orbit ephemeris used for generating the RSLC under a subgroup named "metadata/orbit" and the attitude under a subgroup named "metadata/attitude".

4.4.3.1 Orbit

The orbit ephemeris used for generating the RSLC product can be found under a subgroup named "orbit". This group includes time-tagged antenna phase center position and velocity vectors in Earth Centered Earth Fixed (ECEF) cartesian coordinates. In nominal operations, this would be the Medium Orbit Ephemeris (MOE) state vectors that were used by the L1 processor.

4.4.3.2 Attitude

The attitude state vectors used for generating the RSLC product can be found under a subgroup named "attitude". This group includes time-tagged quaternions and Euler Angles representing the orientation of the radar antenna in the Earth Centered Earth Fixed (ECEF) cartesian system. In nominal operations, this would be the Precise Radar Pointing (PRP) state vectors that were used by the L1 processor [\[RD6\].](#page-9-3)

4.4.4 Geolocation Grid

Section 5.7 contains information describing the radar geometry of the sensor during data taking in the group "/science/LSAR/RSLC/metadata/geolocationGrid". The geolocationGrid cubes are referenced over the radar-grid which is defined by the coordinate vectors slantRange, zeroDopplerTime, and heightAboveEllipsoid. Normals are with respect to the WGS84 ellipsoid.

Geolocation grid cubes also provide the following list of radar geometry information in the associated HDF5 datasets:

- 1. The mapping of the zero-Doppler grid to the geographic grid is described by the cubes datasets "coordinateX" and "coordinateY", expressed in units defined by the EPSG code in "geolocationGrid/epsg".
- 2. The line-of-sight (LOS) unit vector, i.e., the vector from the target to the sensor, is defined by the datasets "losUnitVectorX" and "losUnitVectorY" which contain respectively the east and north components of the LOS unit vector in the east-north-up (ENU) coordinate system. Note that the third component of the LOS unit vector is not provided in the product as it can be simply derived from the other two components as:

$$
losUnitVectorZ = \sqrt{1 - losUnitVectorX^2 - losUnitVectorY^2}
$$

- 3. The along-track unit vector represents the projection of the along-track vector at the ground height. It is defined by the datasets "alongTrackUnitVectorX" and "alongTrackUnitVectorY" containing respectively the east and north components of the along-track unit vector in UTM coordinates.
- 4. The incidence angle, i.e., the angle between the LOS vector and the normal to the ellipsoid at the target height, is given by the dataset "incidenceAngle".
- 5. The elevation angle, defined as the angle between the LOS vector and the normal to the ellipsoid at the sensor, is provided as "elevationAngle".
- 6. The ground track velocity which contains the absolute value of the platform velocity scaled at the target height is given as "groundTrackVelocity".

5 PRODUCT SPECIFICATION

5.1 Dimensions and Shapes

To simplify the description of the layout of data within the HDF5 file, we will use a table of dimensions and shapes to represent the relationship between similarly sized datasets. The entries in this table do not present actual datasets in the HDF5. This table is meant to be a guide to interpreting the shapes of the datasets in subsequent subsections.

Name	Shape	Description
scalar	scalar	scalar values
numberOfDatatakes	scalar	number of datatakes in product
numberOfObservations	scalar	number of observations in product
numberOfFrequencies	scalar	Number of L-SAR frequencies in product
zeroDopplerTimeLength	scalar	Number of lines in all L-SAR imagery datasets
numberOfFrequencyAPolarizations	scalar	Number of polarization layers associated with L-SAR
		frequency A
frequencyASlantRangeWidth	scalar	Number of pixels in all L-SAR frequency A imagery datasets
complexDataFrequencyAShape	(zeroDopplerTimeLength,	Shape associated with L-SAR frequency A imagery datasets
	frequencyASlantRangeWidth)	
numberOfFrequencyBPolarizations	scalar	Number of polarization layers associated with L-SAR
		frequency B
frequencyBSlantRangeWidth	scalar	Number of pixels in all L-SAR frequency B imagery datasets
complexDataFrequencyBShape	(zeroDopplerTimeLength,	Shape associated with L-SAR frequency B imagery datasets
	frequencyBSlantRangeWidth)	
validSamplesShape	(zeroDopplerTimeLength, 2)	Shape associated with L-SAR valid samples dataset
geolocationCubeShape	(geolocationCubeHeight,	Shape associated with metadata cubes
	geolocationCubeLength,	
	geolocationCubeWidth)	
geolocationCubeHeight	scalar	Height dimension of the metadata cube
geolocationCubeLength	scalar	Length dimension of the metadata cube
geolocationCubeWidth	scalar	Width dimension of the metadata cube
dopplerCentroidTimeLength	scalar	Length dimension of Doppler centroid grid
dopplerCentroidSlantRangeWidth	scalar	Length dimension of Doppler centroid grid
dopplerCentroidShape	(dopplerCentroidTimeLength,	Shape of the Doppler centroid grid
	dopplerCentroidSlantRangeWidth)	
calibrationTimeLength	scalar	Length of calibration LUTs
calibrationSlantRangeWidth	scalar	Width of calibration LUTs
calibrationScaleShape	(calibrationTimeLength,	Shape of calibration LUTs
	calibrationSlantRangeWidth)	
antennaPatternComplexShape	(calibrationTimeLength,	Shape of antenna pattern datasets
	calibrationSlantRangeWidth)	
crosstalkComplexShape	scalar	Shape of crosstalk datasets
orbitListLength	scalar	Number of orbit state vectors

Table 5-1 Table of dimensions and shapes in RSLC product

5.2 Product Identification

Table 5-2 NISAR HDF5 variables used for product identification

5.3 Radar Imagery

Table 5-3 NISAR HDF5 variables related to SAR imagery

5.4 Calibration Information

Table 5-4 NISAR HDF5 variables related to calibration

5.5 Processing Information

Table 5-5 NISAR HDF5 variables related to processing parameters

5.6 Other Radar Metadata

Table 5-6 NISAR HDF5 variables related to useful radar metadata

5.7 Geolocation Grid

Table 5-7 NISAR HDF5 variables related to metadata cube

6 METADATA CUBE

In this section, we provide an overview of the metadata cubes used to store spatially-varying ancillary data in the secondary layers of the NISAR L-SAR product HDF5 granules. Note that this sparse representation is to assist users in ingesting and analyzing NISAR products within existing GIS software and is not meant to replace traditional representations of SAR data within the product granules or traditional processing approaches with radar geometry-aware software.

Metadata cubes are represented as three-dimensional arrays in the NISAR product HDF5 modules [\(Figure 6-1\)](#page-53-2). The axes of the array are interpreted as (height, increasing azimuth time, and increasing slant range) in case of radar geometry products and as (height, decreasing northing, and increasing easting) in case of geocoded products. The data is organized with height as the first axis, as this allows one to directly ingest data as GCPs or rasters into existing GIS software. Each height layer is the same size. Metadata cubes will have fixed grid spacing (3 km in azimuth/northing x 1 km in slant range/easting x 1.5 km in height and will allow for easy merging when multiple products along the same imaging track are to be concatenated. The metadata fields on this coarse resolution grid will be evaluated using traditional radar processing approaches without approximations. The metadata cube will also span a field slightly larger than the original image product to allow users to interpolate data without introducing edge effects. Such low-resolution representation of slowly varying parameters has been demonstrated for InSAR products and processing [\[RD5\].](#page-9-4)

Figure 6-1. Metadata cube layer schematic

6.1 Metadata Cube Interpolation Example

We provide here a conceptual example of how these metadata cubes can be used within an existing GIS framework. Let us consider a GUNW product on a UTM Zone 10 grid. We use a geocoded product for the demonstration but the presented approach can be easily extended to radar coordinate products by replacing northing axis by azimuth time and easting axis by slant range.

Suppose we are interested in computing the Perpendicular Baseline (Bperp) at a pixel of interest located at UTM coordinates point (Px,Py). Since these are coordinates on a map domain, we can look up a DEM to get the height at this point. The three-dimensional point of interest then becomes $(Px, Py, h(Px, Py))$.

The metadata cube for Perpendicular baseline can be thought of as a three-dimensional field Bperp(x,y,z) – even though it is oriented as (Nz,Ny,Nx) in the HDF5 file for ease of use with a GIS. The user can use standard built-in regular grid three-dimensional interpolation routines in languages like MATLAB (e.g, interp3), IDL or Python (e.g, RegularGridInterpolator) to interpolate the Bperp array. We recommend cubic interpolation for best results. If a threedimensional interpolator is not available, one could use two-dimensional cubic interpolation for each height layer followed by a one-dimensional cubic interpolation in the following manner:

1. Populate $f(i)$, $i=0,...Nz-1$ by two-dimensional cubic interpolation of each height layer:

$$
f(i) = Bperp\left[i, \frac{Py - Cymax}{Cdy}, \frac{Px - Cxmin}{Cdx}\right]
$$

where the numbers in the square brackets indicate indices into the three-dimensional cube. For example, if we are interested in the point (107590.0 East, 555870.0 North, 300.0 Height), we would interpolate at Row 7.71 and Column 10.59 for each height layer.

2. Interpolate f(i) using one-dimensional cubic interpolation:

$$
Bperp(Px, Py, h(Px, Py)) = f\left[\frac{h(Px, Py) - Czmin}{Cdz}\right]
$$

where the number in the square bracket indicates an index into a one-dimensional array. For example, for a height value of 200.0, we would interpolate at an index of 1.2.

6.2 Metadata Cube Usage Note

Note that the metadata cubes are designed to accommodate one double-precision cube within 1 MB of memory, allowing for information to be easily stored in memory for on-the-fly computation within GIS frameworks or software without much overhead. The metadata cubes are not a replacement for traditional SAR processing approaches or very high-resolution analyses. They are meant to facilitate rapid processing and analysis by non-experts and will serve the needs for most SAR applications. Analyses show that the geolocation error is on the order of 1.5 cm due to interpolation which is significantly smaller than errors from sources such as DEM, orbits, and atmospheric path delay. Interpolation errors for each of the metadata layers will be reported after additional study.

APPENDIX A: ACRONYMS

