



OPERA

**Observational Products
for End-Users from
Remote Sensing Analysis**

Product Validation Document

OPERA

Product Validation Plan

Version 1.5

JPL D-107387 Rev A

September 2, 2022

Paper copies of this document may not be current and should not be relied on for official purposes.

DOCUMENT CHANGE LOG

Revision	Cover Date	Sections Changed	ECR #	Reason, ECR Title, LRS #*
V1.0	4 August 2021	All sections	N/A	First version of the document, LRR058652
V1.1	15 September 2021	All sections	N/A	Second version after table top review, LRR059390
V1.2	6 October 2021	Requirements	N/A	Updated L2 Requirements (Section 2.1); Updated roles and responsibilities (Section 3.6), LRR059912
V1.3	12 November 2021	Title; All sections (headers and references to document as “Product Validation Plan”); Content was changed in Section 2 - 4 (Subsections related to DIST Product) specifically Requirements in 2.1, Field work described in 3.2.3, Responsibilities in 3.6, Validation Strategy in 4.6.	N/A	DIST comments Table-top review for Delta Review; change title to Product Validation Plan, LRR060876
Rev - V1.4	4 February 2022	Responsibilities in 3.6, plan for ongoing validation in 4.7. Signature page and tracking of baselined sections	N/A	Addition of ongoing validations and iterations for baselining, LRR062587
Rev A V1.5	29 July 2022	Section 2.1 L2 Product Requirements; Section 3.2 DSWx Fieldwork; Table 3.4 Summary of Proposed Validation Sites; Figure 3.2 Study site locations; Table 4.1 Summary of Validation sites; Section 4.2 CSLC; section 4.3 RTC; Section 4.5 DSWx Product Validation	N/A	Aligned descriptions of requirements with L2 requirements document. RTC TBDs have been resolved. 3.2 Modified to reflect revised field deployments; Table 3.4 and Figure 3.2 Removed CCNSS and Jones Center, inserted PNWR; Table 4.1 Minor update to cal/val site descriptions. CLSC calval site update. RTC approaches simplified, requirement clarified. 4.5 DSWx Product Validation edited to improve clarity.

* Include the JPL Limited Release System (LRS) clearance number for each revision to be shared with foreign partners.

REV- BASELINED

Chapters/Sections	Signature
<u>All of Section 1 Introduction</u> <u>In Chapter 2 Product Overview</u> 2.2.4 Dynamic Surface Water Extent 2.2.5 Disturbance 2.3 OPERA Product Science Data System <u>In Chapter 3 Validation Overview</u> 3.1 What is Validation Data? 3.2.2 Dynamic Surface Water Extent Field Work 3.2.3 Disturbance Field Work 3.4.1 High-resolution Imagery from Commercial Satellites 3.4.2 Copernicus GLO-30 Digital Elevation Model 3.5 OPERA Validation Science Data System 3.6 Roles and Responsibilities <u>In Chapter 4 Validation Strategies</u> 4.5 Dynamic Surface Water Extent (DSWx) Product 4.6 Disturbance (DIST) Product	Prepared by: Bruce Chapman Concurred by: Heresh Fattahi, Luca Cinquini, Matt Hansen Approved by: David Bekaert, Sean Buckley

REV A BASELINED

Chapters/Sections	Signature
2.1 Product requirements 3.2 Dynamic Surface Water Extent Field Work 4.2 Validation strategies for CSLC 4.3 Validation strategies for RTC 4.5 Dynamic Surface Water Extent (DSWx) Product	see next page

SIGNATURES

Prepared by:

Bruce Chapman, Validation Lead for OPERA

Date

Concurred by:

Heresh Fattahi, JPL ADT Lead for OPERA

Date

Luca Cinquini, JPL SDS Lead for OPERA

Date

Matt Hansen, UMD Lead for OPERA DIST Product

Date

Approved by:

Steven Chan, Project Scientist for OPERA

Date

David Bekaert, Project Manager for OPERA

Date

Key Authors

Bato Grace	<i>Jet Propulsion Laboratory, California Institute of Technology</i>
Bekaert David	<i>Jet Propulsion Laboratory, California Institute of Technology</i>
Chapman Bruce	<i>Jet Propulsion Laboratory, California Institute of Technology</i>
Hansen Matt	<i>University of Maryland</i>
Jones John W	<i>United States Geological Survey</i>
Kim Jin Woo	<i>Southern Methodist University</i>
Logan Thomas	<i>University of Alaska, Fairbanks</i>
Lu Zhong	<i>Southern Methodist University</i>
Marshak Charles	<i>Jet Propulsion Laboratory, California Institute of Technology</i>
Meyer Franz	<i>University of Alaska, Fairbanks</i>
Pickens Amy	<i>University of Maryland</i>
Sangha Simran	<i>Jet Propulsion Laboratory, California Institute of Technology</i>



Part of 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).

TABLE OF CONTENTS

1 Introduction	10
1.1 Background	10
1.2 Scope and Objectives	10
1.3 Document Structure	11
1.4 Applicable Documents	12
2 OPERA Product Overview	13
2.1 Level-2 (L2) Product Requirements	13
2.2 OPERA Data Products	17
2.2.1 Coregistered Single Look Complex (CSLC) Product	19
2.2.2 Radiometric and Terrain Corrected (RTC) Product	19
2.2.3 Displacement (DISP) Product	19
2.2.4 Dynamic Surface Water Extent (DSWx) Product	20
2.2.5 Disturbance (DIST) Product	20
2.3 OPERA Product Science Data System (SDS)	21
3 Validation Overview	21
3.1 What is Validation Data?	21
3.2 OPERA Validation Field Sites	23
3.2.1 Corner Reflectors for CSLC and DISP	23
3.2.1.1 Deployment Timeline	24
3.2.2 Dynamic Surface Water Extent Field Work	26
3.2.2.1 Deployment Timeline	29
3.2.3 Disturbance Field Work	30
3.2.3.1 Deployment Timeline	31
3.3 Resource networks	32
3.3.1 Everglades Depth Estimation Network (EDEN)	32
3.3.2 USGS river and tide gauge network	33
3.3.3 Existing Corner Reflector Arrays	33
3.3.4 UNR, UNAVCO, USGS and JPL SGP GNSS data	34
3.3.5 NISAR wetland Validation data	34
3.4 Ancillary Data Products	35
3.4.1 High-resolution Imagery from Commercial Satellites	36
3.4.1.1 Planet Labs RapidEye and PlanetScope	36
3.4.1.2 MAXAR	36

3.4.2 Copernicus GLO-30 Digital Elevation Model (DEM)	37
3.4.3 Height Above Nearest Drainage	37
3.4.4 Alaska Satellite Facility (ASF) GAMMA Sentinel-1 RTC products	37
3.4.6 Global Forest Change 2000–2019	37
3.4.7 NASA Imagery and Validation Data	38
3.5 OPERA Validation Science Data System (SDS)	38
3.5.1 The Validation Database	39
3.6 Roles and Responsibility for Validation	39
4 Validation Strategies	41
4.1 Preprocessing for Validation	43
4.1.1 Clouds and Cloud Shadows	43
4.1.2 Radio Frequency Interference (RFI)	44
4.1.3 Propagation delays	45
4.1.4 General Image Anomalies	45
4.2 Coregistered SLC product	45
4.3 Radiometric Terrain Correction Product	48
4.4 Displacement Product	53
4.5 Dynamic Surface Water Extent Product	58
4.6 Disturbance Product	63
4.6.1 Validation Strategy for Disturbance Detection in DIST suite	63
4.6.2 Assessment Strategy of the Vegetation Cover Indicator Layer	65
4.7 Ongoing validation of OPERA products	66
4.8 Validation deliverables	67
4.8.1 Validation products	67
4.8.2 Validation reports and peer review	68
References	69
Acronyms	73
Appendix A: Sensors	75
A.1 Sentinel-1 A/B	75
A.2 NISAR	75
A.3 SWOT	75
A.4 Landsat-8	76
A.5 Sentinel-2 A/B	76
Appendix B: Water Classes for the DSWx Product	76

Table of figures

Figure 2-1. OPERA Products dependency and external inputs datasets to OPERA products that are accessed through the NASA Distributed Active Archive Centers (DAACs).

Figure 3.1: Proposed sites of the corner reflectors (triangles) to be deployed by OPERA across San Andreas Fault's creeping segment

Figure 3.2: HLS tile boundaries covering field site regions proposed for Dynamic Surface Water Extent DSWx product Validation.

Figure 4.1: RFI in Sentinel-1 data over the Middle East.

Figure 4.2: Examples of the sequential network of CSLCs, Temporal matching using normalized cross-correlation

Figure 4.3: CSLC with the estimated phase center, Example of an absolute location error analysis

Figure 4.4: Coregistered SLC Validation sites from Table 4.1 displayed geographically

Figure 4.5: Geocoded Sentinel-1B SAR backscatter (VV) without (left) and with (right) radiometric terrain correction (RTC)

Figure 4.6: RTC image and the distribution of backscatter

Figure 4.7: Proposed RTC-SI Validation Sites.

Figure 4.8: Error estimation results for a fault study in the Anatolian region.

Figure 4.9: Proposed Validation displacement sites from Table 4.3 displayed geographically.

Figure 4.10: Temporal coherence map

Figure 4.11 : Locations of Validation sample sites

Figure 4.12: Example of Landsat OLI data, USGS DSWE and RapidEye-based water mask

Figure 4.13 : high resolution optical imagery and open water mask. Landsat imagery coincident with high-resolution image and aggregated/re-projected high resolution water mask into Landsat frame.

Figure 4.14 : Bivariate map indicating sample size and agreement rate with water level as measured at EDEN gauge sites and 'inundation' as detected by a prototype DSWE subpixel inundation detection algorithm.

Figure 4.15: Transect-based data collection using an Accupar LP-80 to generate satellite-product relevant vegetation cover information for inundated Everglades wet prairie

Figure 4.16: Above an example of DIST detection

Table of Tables

Table 2.1: Summary of OPERA Product characteristics.

Table 2.2: Overview of the OPERA product release schedule to the public.

Table 3.1: Summary of validation periods grouped by date.

Table 3.3: Field campaign timeline for corner reflector deployment and maintenance.

Table 3.4: List of DSWx field experiment sites.

Table 3.5: Summary of DSWx Field deployment.

Table 3.7: Summary of Resource Network

Table 3.8: Summary of Validation corner reflector networks.

Table 3.9: Summary of important GNSS networks over North America.

Table 3.10: Summary of ancillary products for Validation.

Table 3.12 Summary of Validation roles and responsibilities.

Table 4.1: Summary of Validation sites for the various OPERA data products.

Table 4.2: List of 12 RTC S1 Validation Sites.

Table 4.3: Proposed Validation displacement sites

Table 4.4: Summary of Validation deliverables for each OPERA product.

1 Introduction

1.1 Background

The JPL OPERA (Observational Products for End-Users from Remote Sensing Analysis) project is funded to develop and implement three products in support of the [Satellite Needs Working Group \(SNWG\)](#) cycle-2 activities: (1) a near-Global Surface Water Extent Product from optical and Synthetic Aperture Radar (SAR) data, (2) a near-Global Land Surface Change Product from optical data, and (3) a North America Land Surface Deformation Product from SAR data. These products respond to needs identified by various federal agencies during the 2018 cycle (a.k.a. cycle-2) of the SNWG. In addition to the above products, there are two intermediate products: (1) a North America Land coregistered Single Look Complex (CSLC) stack product and (2) a near-Global land surface Radiometric Terrain Corrected (RTC) product both derived from SAR data. OPERA's current scope of work provides operational funding until the end of FY 2025, with the various products delivered to and distributed by three NASA Distributed Active Archive Centers (DAACs).

NASA ensures that all data required for Earth system science research are not only archived but include data quality assessments such as data maturity. The NASA Earth Science "Data Maturity Levels" subdivide validated products into [four stages of validation](#):

- **Stage 1 Validation:** Product accuracy is estimated using a small number of independent measurements obtained from selected locations and time periods and ground-truth/field program efforts.
- **Stage 2 Validation:** Product accuracy is estimated over a significant set of locations and time periods by comparison with reference in situ or other suitable reference data. Spatial and temporal consistency of the product and with similar products has been evaluated over globally representative locations and time periods. Results are published in the peer-reviewed literature.
- **Stage 3 Validation:** Product accuracy has been assessed. Uncertainties in the product and its associated structure are well quantified from comparison with reference in situ or other suitable reference data. Uncertainties are characterized in a statistically robust way over multiple locations and time periods representing global conditions. Spatial and temporal consistency of the product and with similar products has been evaluated over globally representative locations and periods. Results are published in the peer-reviewed literature.
- **Stage 4 Validation:** Validation results for stage 3 are systematically updated when new product versions are released and as the time-series expands.

All OPERA products will satisfy Stage 2 validation or higher.

1.2 Scope and Objectives

Validation activities refers to *validating* the resulting OPERA data products. More precisely, per [the NISAR Calibration and Validation document](#):

- *Validation:* The process of assessing by independent means the quality of the data products derived from the system outputs.

The current Product Validation Plan document is focused on the validation of products. Validation can be developed independently from the product derivation using the Level-2 requirements (see Section 2.1 and [AD2]). The requirements provide performance baselines that must be verified using independent data from in situ measurements or ancillary sources. The strategies for performing these validation is the thrust of this document.

There is a project requirement for ongoing validation activities for some products during the 3-year production phase of this project to ensure that the generated products consistently satisfy the L2 requirements. The frequency of this validation will be clarified in future versions and will depend on the OPERA product being validated.

Algorithm verification is within the purview of the OPERA Algorithm Development Team (ADT). As such, verification activities will be covered in the OPERA Algorithm Theoretical Basis Document (ATBD), and may include, if necessary, any calibration of algorithmic parameters. The ATBD will be released after the current version of the Product Validation Plan document. We expect by the next release of the Product Validation Plan and with the upcoming release of the ATBD to have a better roadmap for verification activities which will be communicated both in future versions of this document and the ATBD.

Additionally, systematic Quality Assurance (QA) of the products will be detailed in the ATBD as scoped out in the ADT Level 3 requirements. Relevant QA for Validation activities will be highlighted in future releases of this document as such QA processes will help validation; indeed, OPERA Science and Data Systems will run all the available OPERA QA software to ensure the OPERA products that are used for validation match those that will be generated operationally. There will be additional, sometimes manual, selection and preprocessing of the OPERA products outside of systematic QA that is done for validation activities, and this will be clearly identified in the document (discussed in Section 4.1).

We will refer to the Validation team and the Validation data as they relate to validation activities. As indicated above, all OPERA data products will satisfy Stage 2 Validation or higher according to each product's Level 2 requirements. OPERA will produce continental scale products with minimum accuracy requirements captured in [AD2]. These requirements are further reviewed in section 2.

1.3 Document Structure

The document is organized linearly building toward the description of the validation strategies. The document first provides product descriptions and requirements (Section 2), then elaborates on the ancillary data sources and field work that will be utilized for product validation (Section 3), and finally describes the validation strategy for each product (Section 4). Section 2 and 3 provide the necessary ingredients for understanding the validation strategy discussed in Section 4. Some readers may be interested in first learning the specific validation strategies for the various OPERA products, and in that case, such readers may want to skip first to Section 4, before returning to the previous sections as needed.

Within Section 2, the product requirements from [AD2] and product descriptions from [AD3] are reviewed including the L2 product accuracy requirements needed for the validation of each product. Note that many of the requirements describing the product spatial scale and temporal frequency are omitted here unless they directly impact the validation strategy. Such requirements can be found in the original product requirement document [AD2]. We also briefly discuss the OPERA Product Science Data System (SDS) that will be used to generate OPERA products and the Validation team will access to obtain provisional products for validation.

In Section 3, an overview of the data used for Validation is introduced. Validation data is the independent data sources that are used to validate OPERA products. This includes the in situ measurements made from field measurements and ancillary data sources. Some of OPERA products will require new field measurements for validation. This section describes the field sites, instruments, and the timelines for such field work. The precise validation strategies that will employ this field data will be discussed in Section 4 and readers may wish to read relevant portions where appropriate. The roles and responsibilities of the OPERA Project Science Team (PST) for validation activities are outlined in this section as well. Some validation data will be used to validate multiple OPERA products and we will carefully record which product validations utilize which ancillary data products and/or field data. We also introduce the Validation Database that will provide scientists easy application protocol interfaces (APIs) to ingest, store, and access the Validation data.

Section 4 is what the document builds towards: elaborating upon the precise validation strategies for each OPERA product. Section 4 details how each OPERA product is shown to satisfy its L2 requirement using the field data, ancillary data, resource networks, etc. Some readers may wish to read (or skim) Section 4 first to better understand the validation strategies. This may provide context of the entire document.

Some validation activities are still under active development. In other words, there are product validation strategies that are still being formulated, particularly for OPERA products with release dates occurring later in the project timeline (see Table 3.1). Currently, the Dynamic Water Extent for HLS data (DSWx-HLS) and the Disturbance Product (DIST) product are the products that have the earliest Validation release and therefore will have the most mature validation strategies. Areas that require future augmentation in subsequent versions will be clearly identified. Additionally, the Algorithm Theoretical Basis Document (ATBD) has not been published and the precise algorithmic details for each product can only be referenced using the source material identified in [AD3]. Each product algorithm will have systematic Quality Assurance (QA) that will be detailed in the ATBD and further benefit the validation activities. Any OPERA specific implementations from the ATBD that will require additional validation activities will be identified and explained in later versions of this document. Fortunately, the validation of the OPERA products can be described almost entirely using the product requirements and the product description.

1.4 Applicable Documents

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

- AD1: SNWG Program-Level Requirements document (Version August 2, 2020)
- AD2: OPERA Project requirements document (JPL D-107391)
- AD3: OPERA product description document (JPL D-107389)

Additional reference documents are also provided as a bibliography and cited throughout the document for more detailed information about algorithms, input datasets, etc.

2 OPERA Product Overview

This section briefly reviews the OPERA products and requirements. First, the product requirements (2.1) from [AD2] are summarized. These requirements specify a performance baseline that must be satisfied for each OPERA product. The precise strategy for verifying these requirements will be detailed in Section 4. In other words, the requirements are “*what*” must be verified. The “*how*” is the topic of validation strategies in Section 4.

Also, in Section 2.2, we review the product descriptions from [AD3] so that readers less familiar with OPERA science data products can understand the high level scope of each product. In Section 2.3, we introduce the OPERA Product Science Data System (SDS) that is used to generate the OPERA products.

2.1 Level-2 (L2) Product Requirements

This section describes the Level-2 (L2) product requirements from [AD2] and provides a statement of *what* needs to be validated. These requirements are translated into actionable strategies in Section 4, demonstrating *how* these requirements will be satisfied using Validation data from Section 3.

The OPERA Level-1 requirements [AD1] were translated into Level-2 system requirements [AD2] on the various subsystems, including those specifically related to Level-2 and Level-3 products. These L2 requirements fall into four general categories: (a) resolution requirements, (b) radiometric and spatial location accuracy requirements, (c) product accuracy requirements and (d) latency and throughput requirements. These are summarized in the level 2 data product tables [AD2]. The temporal sampling of each product may be found in table 2.1. The accuracy requirements will be validated using the proposed strategies in Section 4.

Some elements of the requirements may state that they are To Be Confirmed (TBC) or To Be Determined (TBD). Each element so labeled either has a specific timetable for resolution (TBC) or an associated study (TBD). We have highlighted where such changes may be expected using the footnotes. Additionally, some requirements will have language that requires explanation or clarification. This language is addressed in the notes wherever possible.

Coregistered Single Look Complex (CSLC) Product

Accuracy requirement	Notes
The Sentinel-1-based CSLC product (CSLC-S1) shall have a Relative Location Error of 0.5 meters and 0.75 meters or better in ground range and azimuth directions respectively, in areas with high cross-correlation accuracy, excluding the effects of DEM errors and ionospheric scintillation on SAR acquisitions, in at least 80% of all validation products considered. The product's Relative Location Error shall be determined relative to a single reference image within the stack. The Sentinel-1-based CSLC product (CSLC-S1) shall have an Absolute Location Error 1.5 meters or better in both the ground range and azimuth directions, excluding the effects of DEM errors and ionospheric scintillation on SAR acquisitions, in at least 80% of all validation products considered.	Currently, the errors are reported in range-doppler coordinates, which can be converted to map coordinates using standard transformations related to the SAR viewing geometry. Depending on whether the final CSLC product is geocoded will determine how the errors are captured in the final language of this requirement. The CSLC geolocation errors are driven by interferometry needs and sensor imaging modes. The extra Doppler due to Sentinel-1 beam rotation in TOPS mode imposes more precise azimuth geolocation accuracy to avoid interferometric phase jumps at burst boundaries. This does not apply to NISAR data acquired in sweepSAR mode. “high cross-correlation accuracy”: in the absence of decorrelated pixels due to vegetation, snowfall, flood, catastrophic damage and other phenomena which does not allow reliable estimation of relative shifts between acquisitions using data driven methods.
The NISAR-based CSLC product (CSLC-NI) shall have a relative geolocation accuracy of 0.7 m or better in slant range and azimuth directions over pixels with interferometric coherence of at least 0.5 (in the absence of decorrelated pixels due to vegetation, snowfall, flood, catastrophic damage and other phenomena which does not allow reliable estimation of relative shifts between acquisitions using data driven methods) for at least 80% of all validation products considered. The product's relative accuracy shall be made relative to a single reference image within the stack. The NISAR-based CSLC product (CSLC-NI) shall have an absolute geolocation accuracy of 5 meters or better [TBC] in the slant range direction, and 10 meters or better [TBC] in the azimuth direction over bare-ground areas with slopes less than 1 m per 30 m, excluding the effects of DEM errors for at least 80% of all validation products considered except when SAR acquisitions are affected by ionospheric scintillation.	Measured with respect to a reference image in the stack. No additional requirement will be set for these decorrelated pixels.

Radiometric Terrain Corrected (RTC) Product

Accuracy requirement	Notes
The Sentinel-1-based RTC product (RTC-S1) shall meet an absolute and relative geolocation accuracy better than or equal to 6 meters given the 30 meter RTC-S1 product resolution (i.e. 20% of the product resolution), excluding the effects of DEM errors, for at least 80% of all validation products considered. The median radar backscatter of the OPERA RTC-S1 product over an area of foreslope shall be within 1dB of the median radar backscatter over an area of backslope for forested land-types, for at least 80% of all validation products considered.	Relative accuracy is relative to product time series. For this requirement, forested land types are defined as classes ABC-XYZ based on 2016-2019 CGLC-100 land cover classification.

Displacement (DISP) Product Suite

Accuracy requirement	Notes
The Sentinel-1 DISP product (DISP-S1) shall measure from a single line-of-sight geometry surface displacement rates with an uncertainty of 3 mm [TBC] per yr or better over length scales within $0.1 \text{ km} < L < 50 \text{ km}$ from 4 years of regularly sampled (at least 80% of Sentinel-1 acquisitions at 12 days sampling or better) Sentinel-1 A/B data (VV polarization imagery in IW-mode), in regions where the interferometric signal is maintained (i.e., the temporal coherence is above 0.5).	“Temporal coherence above 0.5” is equivalent to “the interferometric coherence throughout the related time series is above 0.5, when interferometric coherence is computed across sequential pairs”.
The NISAR DISP product (DISP-NI) shall measure from a single line-of-sight geometry surface displacement rates with an uncertainty of 5 mm [TBC] per yr or better over length scales $0.1 \text{ km} < L < 50 \text{ km}$ from 1 year of regularly sampled (i.e. at least 80% of acquisitions from nominal 12 days repeat passes) NISAR data (i.e. from HH polarization imagery from the L-band main band), in regions where the interferometric signal is maintained (i.e., the temporal coherence is above 0.5).	NISAR L-band radar is capable of acquiring data on lower and upper parts of a total 80 MHz range bandwidth called main and side bands respectively. The OPERA project will produce displacement products using main-band data. “Temporal coherence above 0.5” is equivalent to “the interferometric coherence throughout the related time series is above 0.5, when interferometric coherence is computed across sequential pairs”.

Dynamic Surface Water Extent (DSWx) Product Suite

Accuracy requirement	Notes
Detectable water features shall be 3 hectares in size or larger and at least 200 m in width. The OPERA optical water extent product (DSWx-HLS) shall detect open water with a classification accuracy of at least 80% for at least 80% of all validation products considered. The OPERA optical water extent product (DSWx-HLS) shall detect partial surface water with a classification accuracy of at least 70% and where retrievals are possible.	The water classes used in the below requirements are defined in Appendix B including <i>open water</i>, <i>partial surface water</i>, and <i>water under vegetation</i>. Please refer to the appendix as needed. Water excludes ice and non-liquid states. The classifier that “detects” in this requirement is binary: it detects either “open water” or “not open water”. If the product specifies more classes, the requirement will be applicable to the related binary classifier that can be derived from the more complex one, i.e. combining all classes that are not “open water” into one class.
Detectable water features shall be 3 hectares in size or larger and at least 200 m in width. The OPERA DSWx-S1 shall detect open water with a classification accuracy of at least 80% for at least 80% of all validation products considered. [TBD] The OPERA DSWx-S1 shall detect inundated vegetation in wetlands with a [TBD] classification accuracy (i.e. at a level that will be established given future algorithm and performance testing). [TBD if this applies to more than just wetlands.]	This classification accuracy is the total accuracy over the pixels that are considered for validation. The retrieval of partial surface water will be primarily limited by vegetation density. While validation data can be generated to assess this limitation, no suitable globally applicable dataset is currently available to designate where these conditions occur.
Detectable water features shall be 3 hectares in size or larger and at least 200 m in width. The OPERA DSWx-NI shall detect open water with a classification accuracy of at least 80% for at least 80% of all validation products considered. The OPERA DSWx-NI shall detect inundated vegetation in wetlands with a 80% classification accuracy (i.e. at a level that will be established given future algorithm and performance testing). [TBD if this applies to more than just wetlands.]	The retrieval of inundated vegetation will be primarily limited by vegetation density and structure, which dictate backscatter returns. While validation data can be generated to assess this limitation, no suitable globally applicable dataset is currently available to designate where these conditions occur. For SWOT and Sentinel-1, there will be a trade study performed by the Algorithm Development Team to determine feasibility of the inundate vegetation detection and will be updated in future versions of the document.
Detectable water features shall be 3 hectares [TBC] in size or larger and at least 200 m [TBC] in width. The OPERA DSWx-SW shall detect open water with a classification accuracy of at least 80% for at least 80% of all validation products considered. [TBD] The OPERA DSWx-SW shall detect inundated vegetation in wetlands with a [TBD] classification accuracy (i.e. at a level that will be established given future algorithm and performance testing).	

Disturbance (DIST) Product Suite

Accuracy requirement	Notes
The disturbance (DIST) products, derived from HLS data, shall identify pixels having at least 50% vegetative cover loss per HLS pixel (approximately 30 m in size) between a current single-date HLS observation as compared to a HLS composite from previous years. The Disturbance Alert (DIST_ALERT) product shall measure disturbances with 80% accuracy per validation pixel and the aggregated annual products (DIST_ANN) shall achieve 90% accuracy per validation pixel for at least 80% of all validation products considered.	Cover (as in vegetation cover) is the amount of skylight orthogonal to the surface that is intercepted by the trait of interest. Vegetation includes woody and herbaceous (aka non-woody) vegetation as is done for the MODIS Vegetation Continuous Field (VCF) product. Loss within an area/pixel over a given time frame is defined to be either 50% vegetation cover decrease. The DIST_ALERT product will validate both confirmed and provisional disturbances. The DIST_ANN product will validate only confirmed disturbances. The accuracy is expected to increase for DIST_ANN because confirmed disturbances, which by definition, have been observed in multiple HLS scenes.

2.2 OPERA Data Products

A detailed overview of the OPERA data products is found in the OPERA product description document [AD3]. Figure 2.1 and Table 2.1 (copied from the product description document) provide a top-level overview of the spatial-temporal resolution and extent, and the data flow between products. OPERA products are referred throughout the document with their short product names. The naming convention as established by the OPERA project is <OPERA product name>-<Input Sensor Acronym>. For example, the Coregistered Single Look Complex (CSLC) product for Sentinel-1 (S1) is called CSLC-S1.

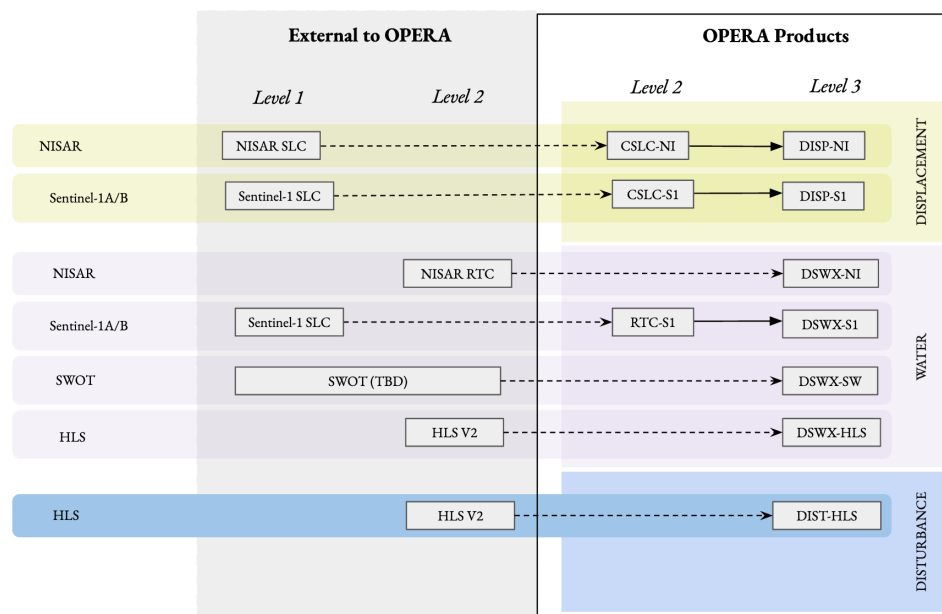


Figure 2.1: OPERA Products dependency and external inputs datasets to OPERA products that are accessed through the NASA Distributed Active Archive Centers (DAACs).

Product name	Coverage	Sensor	Acronyms	Spatial Resolution	Sensor ¹ Temporal Sampling***	Beginning of Product Record****
<i>Coregistered Single Look Complex (CLSC)</i>	<i>North America*</i>	<i>Sentinel-1 A/B</i>	<i>CSLC-S1</i>	~15 m x 5 m (az x range)	12 day (single sensor + same geometry)† 6 day (constellation + same geometry)†	Based on Sentinel-1 A/B Availability (April 3, 2014 - Sentinel-1A launch) (April 25, 2015 - Sentinel-1B launch)
		<i>NISAR</i>	<i>CLSC-NI</i>	~5 m x 3.1/6.25 m (az x range)	12 day (same geometry)†	Based on NISAR Availability
<i>Radiometric Terrain Correction (RTC)</i>	<i>Near-Global**</i>	<i>Sentinel-1 A/B</i>	<i>RTC-S1</i>	30 m	12 day (single sensor + same geometry)† 6 day (constellation + same geometry)†	From end of product validation: December 15, 2022
<i>Displacement Product (DISP)</i>	<i>North America*</i>	<i>Sentinel-1 A/B</i>	<i>DISP-S1</i>	30 m or ~15 m x 5 m (az x range)	12 day (single sensor + same geometry)† 6 day (constellation + same geometry)†	Based on Sentinel-1 A/B Availability (April 3, 2014 - Sentinel-1A launch) (April 25, 2015 - Sentinel-1B launch)
		<i>NISAR</i>	<i>DISP-NI</i>	30 m or ~5 m x 3.1/6.25 m (az x range)	12 day (same geometry)†	Based on NISAR Availability
<i>Disturbance Product (DIST)</i>	<i>Near-Global**</i>	<i>Landsat-8 & Sentinel-2 A/B (HLS)</i>	<i>DIST_ALERT</i>	30 m	16 days for Landsat 8 (same geometry) 10 days for Sentinel-2 (same geometry) ~5 days for Sentinel-2 (constellation + same geometry)	From end of product validation: February 15, 2023
			<i>DIST_ANN</i>	30 m	median average 2.9 days for S2A/B+L8 [Lian, Roy 2017]	
<i>Dynamic Surface Water Extent (DSW_x)</i>	<i>Near-Global**</i>	<i>Landsat-8 & Sentinel-2 A/B (HLS)</i>	<i>DSW_x-HLS</i>	30 m	16 days for Landsat 8 (same geometry)† 10 days for Sentinel-2 (single sensor+same geometry)† ~5 days for Sentinel-2 (constellation + same geometry)† median average 2.9 days for S2A/B+L8 [Lian, Roy 2017]	From end of product validation: February 15, 2023
				30 m	12 day (single sensor + same geometry)† 6 day (constellation + same geometry)†	From end of product validation: June 16, 2023
		<i>NISAR</i>	<i>DSW_x-NI</i>	30 m	12 days (same geometry)†	From end of product validation: February 6, 2024
		<i>SWOT</i>	<i>DSW_x-SWOT</i>	TBD (100 m or 120 m)	21 days (same geometry)†	From end of product validation: January 27, 2025

Table 2.1: Summary of OPERA Product characteristics. * = USA and US Territories within 200 km from US border, Canada, and all mainland countries from the southern US border up to and including Panama. ** = all landmasses excluding Antarctica. *** = Subjected to sensor availability, ****This indicates the beginning of the product record, not the beginning of OPERA processing. † = Descending or Ascending. ¹Temporal sample of a pixel on the ground for a given product requires sensor ground simulations (TBD), hence sensor sampling is provided.

Table 2.2: Overview of the OPERA product release schedule to the public.

Release	Products	Product Release
1	<i>Optical* DSW_x</i> <i>Optical* DIST</i>	<i>End of April 2023</i> <i>Mid of February 2023 (provisional)</i> <i>End of September 2023</i>
2	<i>S1A/B RTC</i> <i>S1A/B CSLC</i>	<i>End of September 2023</i> <i>End of September 2023</i>
3	<i>NISAR Coreg SLC</i> <i>S1A/B DSW_x</i> <i>NISAR DSW_x</i>	<i>End of May 2024**</i> <i>End of May 2024</i> <i>End of May 2024**</i>
4	<i>S1A/B DISP</i>	<i>End of November 2024</i>
5	<i>NISAR DISP</i> <i>SWOT DSW_x</i>	<i>Beginning of July 2025**</i> <i>Beginning of July 2025</i>

**Optical includes both Landsat-8 & Sentinel-2A/B sensors. The optical sensors deliver one product because they use HLS as inputs. All releases are of validated products unless specified.*

*** Subject to change due to changes in the NISAR Launch date*

2.2.1 Coregistered Single Look Complex (CSLC) Product

The Co-registered Single Look Complex (CLSC) product will be a stack of geocoded SLC images for Sentinel-1 (CSLC-S1) and NISAR (CSLC-NI). The CSLC contains individual raster layers representing complex signal return for each of the input polarizations.

2.2.2 Radiometric and Terrain Corrected (RTC) Product

The Radiometric Terrain Correction (RTC) product is the γ^0 (gamma naught) polarimetric imagery that has been corrected for terrain and SAR geometry according to [David Small, 2011, Shiroma, 2020]. Within OPERA, RTC products are generated for acquisitions of the Sentinel-1 mission (RTC-S1). The RTC products are distributed as global tiles aligned with Harmonized Landsat and Sentinel-2 (HLS) data.

2.2.3 Displacement (DISP) Product

The Displacement (DISP) product is derived from an interferometric time-series analysis of the co-pol Coregistered Single Look Complex products (CSLCs). They capture the displacement in the radar line-of-sight. The Sentinel-1 C-band and NISAR L-band DISP products are generated separately and are called DISP-S1 and DISP-NI, respectively and are created for each acquisition.

2.2.4 Dynamic Surface Water Extent (DSWx) Product

The Dynamic Surface Water eXtent (DSWX) products map pixel-wise surface water detections using optical or SAR imagery. Each input dataset will produce distinct, but complimentary products: DSWX-HLS, DSWX-S1, DSX-NI, and DSWX-SWOT. The DSWx tiles will be aligned with HLS data. The algorithm for the DSWx-HLS product is a revision of the Dynamic Surface Water Extent (DSWE) algorithm found in [Jones 2019] and the related USGS DSWE product [USGS link]. The DSWx algorithm for NISAR, Sentinel-1, and SWOT are being formulated and the performance baselines will need to be determined in the next version of this document. As algorithms are developed during the OPERA project, the requirements will be clarified and validation activities will be described in future versions of this document.

There are 4 relevant classes: no-water, open water, partial surface water, and inundated vegetation. See Appendix B for an expanded discussion.

2.2.5 Disturbance (DIST) Product

The land Disturbance product (DIST) maps per pixel vegetation disturbance (specifically, vegetation cover loss) from the Harmonized Landsat-8 and Sentinel-2 A/B (HLS) scenes. Vegetation disturbance is determined when there is a decrease in vegetation cover within an HLS pixel. The product also will provide auxiliary generic disturbance information as determined from the variations of the reflectance through the HLS scenes to provide information about more general disturbance trends. In addition, the DIST product contains intermediate layers associated with vegetation cover including a historical vegetation cover determined from previous calendar years and an anomaly value (which is the difference between the historical vegetation and the current vegetation cover estimate). From these two layers, the current vegetation cover estimate can be obtained. All these layers are not formally validated with respect to project requirements, they are used by the disturbance layers and are verified using additional activities detailed in Section 4.

Two DIST products are generated within the DIST product suite: a) the DIST_ALERT product capturing vegetation disturbance at the cadence of HLS sampling (median average 2.9 days for S2A/B+L8 [Li and Roy, 2017]) and b) the DIST_ANN product capturing the confirmed changes of the DIST_ALERT products from previous calendar year. The date of the first disturbance, the historical vegetation cover, and current vegetation cover estimate are included in both products (for the annual product, the vegetation cover in the respective year is the maximum vegetation cover over the relevant period). Clearly, the primary difference between these two DIST products is their temporal sampling. Each DIST_ALERT product is associated with a unique HLS scene and is used to track vegetation disturbances at the temporal frequency of the input HLS dataset. The DIST_ANN tracks changes at the annual scale, aggregating changes identified in the DIST_ALERT product. The DIST_ALERT product tracks two categories of changes: *provisional* and *confirmed*. *Provisional changes* are when a pixel is first identified to have vegetation cover loss and *confirmed changes* are when this disturbance is identified consistently through some number of subsequent acquisitions in time. The precise number of scenes required for a *confirmed* status will be determined during the algorithm verification and detailed in the ATBD. If a provisional change is not observed in subsequent images, then this pixel's vegetation cover will continue to be analyzed for future vegetation cover losses. In DIST_ANN, only *confirmed changes* from the associated calendar year are tracked.

2.3 OPERA Product Science Data System (SDS)

The OPERA Product Science Data System (SDS) refers to the *physical* computer infrastructure that will generate all of the above OPERA products. The OPERA SDS consists of two elements: a JPL and a UMD-GLAD element. All products except for the DIST product suite will be generated using the JPL element. The JPL SDS element is based off the NISAR SDS, which is discussed in the [Mission Science Handbook](#) [Accessed Jun 8, 2022] and thoroughly specified in the [online Software Design Document](#) [Accessed Sep 1, 2021]. The UMD-GLAD element is built off the existing Global Land Analysis and Discovery (GLAD) laboratory compute infrastructure at the University of Maryland.

Details of the *design* of the OPERA Product SDS are discussed in the not-yet released OPERA SDS document. It will be referenced in future versions of this Product Validation Plan. A *clone* (or a *re-initiation*) of the OPERA Product SDS will be deployed for the validation team using the same design. This clone will generate products strictly for the validation activities (i.e. by request and coordination with the SDS). Although the physical deployments to generate the nominal OPERA products (OPERA Product SDS) and its clone used to generate provisional products for validation are distinct computer infrastructures, the utilized algorithms, ancillary datasets, and deployed resources will be identical for both systems. More details are discussed in Section 3.5. The reason for this distinct physical deployment for validation, albeit identical design, is to reduce the risk that requests to generate provisional OPERA products do not burden the nominal OPERA Product SDS.

In Section 3.5, we will discuss the OPERA Validation SDS that will store the independent validation data and describe the physical computer infrastructure to assess them. This includes the clone of the product system identified above.

3 Validation Overview

3.1 What is Validation Data?

Validation data refers to the data that can be used to validate OPERA products and are generated or collected independently from the OPERA product generation pipeline. Validation data includes in-situ measurements, ancillary datasets, and data from various resource networks. Some OPERA products will require new field measurements for validation and this section outlines the proposed scope of field work and the related deployment timeline. The validation strategy that utilizes the proposed field work will be introduced in Section 4 and readers are encouraged to consult relevant parts of the section to understand the validation strategy. Conditions regarding areas used for validation that are imposed by some requirements are also discussed in Section 4 (summarized in Table 4.1). Certain Validation data is *not* directly compared to OPERA products, but used to support validation activities. For example, a Digital Elevation Model (DEM) will be used to determine areas of “high slope” that will be used in the validation activities of the RTC-S1. In this section, we will provide only brief descriptions of how Validation data will be used; readers will be given a more comprehensive picture in Section 4.

The Algorithm Theoretical Basis Document (ATBD) for each product will describe the implementation of the algorithm, including its verification. The verification of previously implemented algorithms may require some comparison with other products, including validation products described in this document, to ensure that the algorithms are creating products with expected characteristics based on their previous implementations. If any “calibration” of algorithmic parameters is required, it will be described in detail in each product's ATBD. In this context, “calibration” will refer to the set of operations that establish,

under specified conditions, the relationship between sets of values of quantities indicated by a measuring instrument or measuring system and the corresponding values realized by standards.

In this section, we also introduce the OPERA Validation Database that will support validation activities. When possible, all the Validation Data and provisional OPERA products will be made available to the public. The only caveat is that certain source data, e.g. licensed imagery, cannot be shared with the public due to copyright law. Intermediate datasets derived from the licensed data can be shared so long as they cannot be used to reverse engineer the original licensed dataset. For example, it is permissible to distribute a water mask derived from licensed high-resolution imagery because the image's reflectance values cannot be determined from the classification, i.e. classification maps from reflectance imagery are not one-to-one, and hence the original dataset is not retrievable.

Because some of the Validation data sources discussed below are used to validate multiple OPERA products, we highlight which of the Validation data products will be used to validate which of the OPERA science data products. This section captures the relevant details of the Validation data: its collection, organization, storage, etc. Some readers may wish to skip to Section 4 to understand the precise validation strategies first and return to this section afterwards or as needed.

In section 3.2, the OPERA Validation field sites are introduced, including a brief description of proposed in situ observations, related instrumentation, and the timelines for such work. The OPERA Validation field data is being collected *specifically to validate and calibrate OPERA products*. In Sections 3.3 and 3.4, the resource networks and ancillary datasets are detailed. In Section 3.5, the OPERA Validation Science Data System (SDS) is described. The SDS encompasses the underlying computer infrastructure that will be used to support Validation activities. Of particular import is the Validation database, which will be used to store and retrieve relevant in situ measurements and ancillary data for validation activities. In section 3.7, the roles and responsibilities for the execution of the Validation activities are outlined.

Timelines for Validation for each product are summarized in Table 3.1. The products are grouped according to the five validation periods obtained from the OPERA project schedule. Note there will be periodic validation occurring after the end-date in Table 3.1 depending on the requirements of the OPERA product, the additional temporal sampling required for validation, and the overall workforce required. These considerations will be determined in subsequent Product Validation Plans. We highlight that validation activities for the DSWx suite are spread across the 3 different validation periods. As noted in Table 2.2., DSWx-HLS will be released first because it is the most mature. As noted earlier, the algorithm is adapted from [Jones, 2019], which has already been used to generate North American water extents from Landsat [USGS DOI]. The HLS product will be an extension of this North American water product to the near-global scale. The DSWx algorithm for NISAR, Sentinel-1, and SWOT, in contrast, are being formulated and the performance baselines will need to be determined in the next version of this document. As algorithms are developed during the OPERA project, the requirements will be clarified and validation activities will be described in future versions of this document.

OPERA Product Acronym	Begin Validation Period	End Validation Period
DSWx-HLS	August 2022	January 2023
DIST suite	October 2022 (period 1) February 2023 (period 2)	February 2023 (period 1) September 2023 (period 2)

RTC-S1 CSLC-S1	April 2023	August 2023
CSLC-NI DSW _x -S1 DSW _x -NI	January 2024	June 2024
DISP-S1	April 2024	September 2024
DSW _x -SWOT DISP-NI	January 2025	June 2025

Table 3.1: Summary of validation periods grouped by date. Products (and associated algorithms) with earlier validation periods are more mature than those with later validation periods.

3.2 OPERA Validation Field Sites

In this section, we describe the sites that will require field deployments and additional instruments/sensors specifically for OPERA Validation activities. Relevant field sites and instruments are also introduced. More specific details regarding how the sites integrate into each product’s specific Validation strategy is discussed in Section 4. We note that not all products require field deployments, e.g. the RTC product.

3.2.1 Corner Reflectors for CSLC and DISP

We are deploying at least 4 trihedral corner reflectors (CRs) by Q4 of FY2022 in California (i.e., across the creeping segment of the San Andreas fault (SAF) and in a subsiding region in Central Valley) with different spatial baselines (i.e. between 9 and 25 km) suitable for both L- and C-bands. These will be used for validation of the Displacement product and the CSLC. The Displacement product is the primary impetus for the deployment of these corner reflectors and the precise strategy and location is still being developed. The CSLC validation will be used in conjunction with the numerous corner reflector array networks that are currently already available (see 3.3.3) and standard practice will be employed to verify absolute geolocation accuracy with corner reflectors as in [\[Gisinger et al. 2020\]](#). The details of the CSLC validation strategy as it relates to OPERA is discussed in Section 4.2. The details of the existing DISP validation strategy is discussed in Section 4.5.

Aside from supporting the CSLC validation activities, the proposed configuration of the CR sites would allow multiple experiments to be performed in order to complement the validation of the OPERA DISP products as well as to evaluate the performance of the algorithms that will be used to generate these products. We have consulted with the NISAR Solid Earth team to be synergistic with their Validation plan and collected feedback from scientists across different fields of expertise that will benefit from this new corner reflector deployment. The proposed sites along the SAF are currently chosen so that: 1) they fall within common NISAR and Sentinel-1 footprints, 2) three sites are collinear and perpendicular to the fault trace between Bitterwater, CA and Parkfield, CA, 4) one of the target sites is near the subsidence bowl in Central Valley, CA, 4) are located at different terrains and vegetation cover, and 5) can be co-

located (e.g. 20 - 40 m baseline separation [[Garthwaite et al., 2015](#)]) with an active continuous GNSS (cGNSS) station.

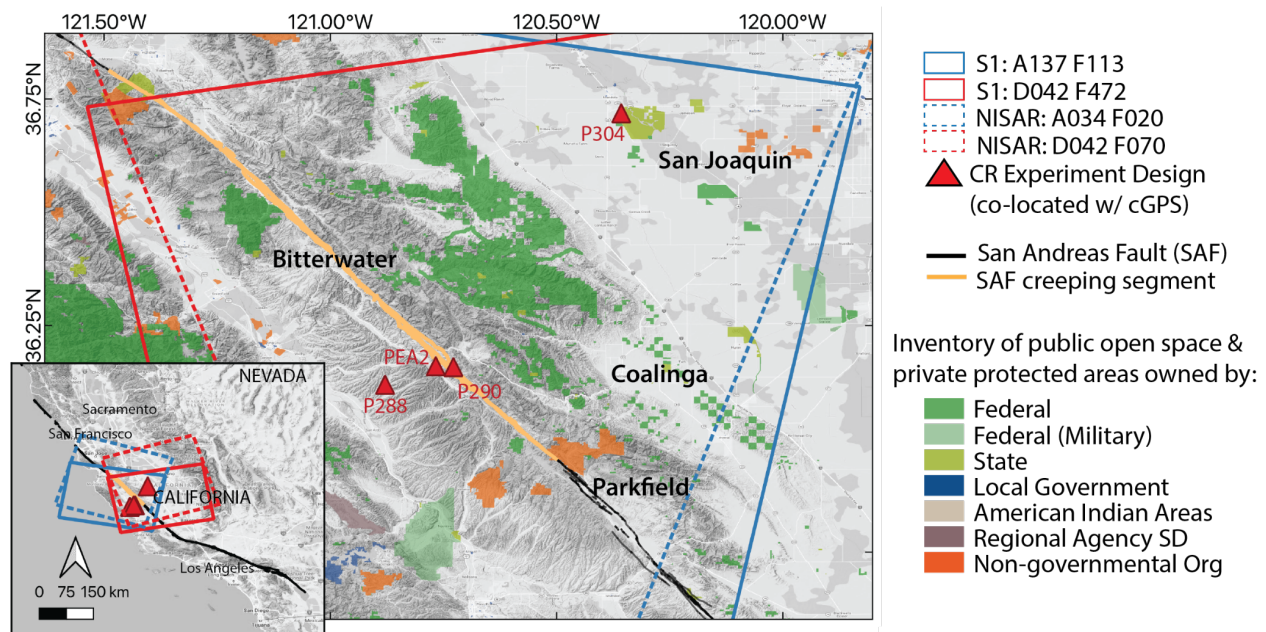


Figure 3.1: Proposed sites of the corner reflectors (triangles) to be deployed by OPERA across San Andreas Fault’s creeping segment with different spatial baselines (i.e. 9 - 25 km) and in a subsiding region in Central Valley, CA. Inset map shows the Sentinel-1 (solid lines) and planned NISAR (broken lines) swaths for both their ascending (blue) and descending (red) tracks. The [fault trace](#) and the inventory of public open space and the [Protected Areas Database](#) are openly available from USGS repositories.

3.2.1.1 Deployment Timeline

We are targeting the deployment of the corner reflectors by Q4 of FY 2022. The deployment needs to meet this aggressive timeline to keep up with the timeline set forth in the larger OPERA schedule.

To ensure a successful deployment, we will: 1) design a validation strategy for the DISP product using the new CRs, 2) select the CR design appropriate for L-band and C-band SAR and related surface deformation, 3) perform standard site reconnaissance, 3) procure the CRs, and 4) complete necessary permitting requests. A summary of the activities can be found in Table 3.3 and the validation activities for the relevant OPERA products are shown as well: DISP-NI, DISP-S1, CSLC-S1, and CSCL-NI. Note that the fieldwork must occur prior to the validation activities. From Table 3.3., we see that only the maintenance and survey of the deployed CR occur during the validation period of the products. The precise validation activities are discussed in Section 4.

For the design of the validation strategy of the DISP product, we are planning the exact locations and orientations of the CRs across SAF and in a subsiding region in Central Valley, CA, to complement the validation of the OPERA DISP products and help evaluate the performance of the algorithms that will be used to generate these products. The proposed CR configuration will allow different experiments to be conducted such as: 1) phase closure assessment experiment, 2) PS-DS experiment over vegetated region, and 3) intercomparison between NISAR (L-band, left-looking) and Sentinel-1 (C-band, right looking) displacements.

While formulating the specifics of the validation strategy, we will also proceed with site reconnaissance activities and the selection of the CR design in the first half of FY 2022. Site reconnaissance will determine the feasibility of the CR locations including whether lease agreements are needed for the deployment, regular access and maintenance by the OPERA project, etc. Note that we plan to select sites that are co-located with or to be in close proximity to cGNSS stations managed by the university consortium UNAVCO (i.e. within a 40-meter radius). Selecting sites in this fashion will help to avoid large delays in permitting requests as such deployments are working with existing field data collection. Specifically, the request to obtain permits for installing the CRs will be coordinated first with UNAVCO. Although we do not anticipate any large delays, the permitting request schedule is planned for two quarters (i.e. Q2 and Q3 of FY 2022) to accommodate any unexpected setbacks. The procurement of CRs will be coordinated with similar current and ongoing acquisitions by the NISAR project, and the schedule to acquire will happen as soon as the selection of the CR is finalized.

For the selection of the CR design, we will work with the NISAR and the UAVSAR projects at JPL ascertaining the appropriate dimensions of the CRs and their suitability for both L-band and C-band Validation activities.

After the CRs have been deployed, we will survey, inspect, and perform maintenance one- to two-times a year starting FY 2023 and onwards.

Activity	FY21	FY22				FY23				FY24				FY25			
CSLC	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Experiment Design	x	x															
Design of CRs		x															
CRs Procurement			x	x	x												
Site reconnaissance fieldwork					x	x											
Permit request					x	x											
CRs Deployment						x	x										
Inspection, maintenance, and survey of CRs								x		x		x		x		x	

Table 3.3: Field campaign timeline for corner reflector deployment and maintenance. The red cells reference validation period from Table 3.1 for CSLS-S1 and the green for CSLC-NI. Similarly, the validation period for DISP-S1 and DISP-NI are shown in light blue and light orange. The bulk of the field work must occur before the validation activities occur. Above, note the maintenance and survey of the CR deployments occur during the validation period as expected.

3.2.2 Dynamic Surface Water Extent Field Work

The DSWx field work is motivated by requirements from Section 2.1 related to the retrievals of the partial surface water and inundated vegetation¹, though data produced from this field work will be used for all validation activities. Currently, there is no suitable applicable dataset to validate partial surface water and inundated vegetation with 30 meter resolution at the near-global scale that OPERA requires. Therefore, OPERA is tasked to provide high quality Validation data from field work to evaluate the DSWx product suite, particularly detection of partial open surface water and inundation under vegetation, and specify performance baselines relevant to these detections.

The detection of partial surface water and inundated vegetation will be limited by vegetation density and input sensor sensitivities. For example, L-band backscatter data over inundated vegetation in wetlands will likely be dominated by double bounce and volumetric scattering and therefore can be differentiated from open water areas, which we expect to have specular returns [Mahdianpari et al. 2017]. Therefore, in addition to mapping areas with partial open surface water and inundation under vegetation, field work will also estimate vegetation density to clarify the performance baseline of the detection of such areas across all the OPERA datasets.

Because no applicable datasets are available, field work for validation activities is crucial. Even with very high-resolution optical imagery (e.g. smaller spatial scale than HLS or SAR input data), identifying inundation between and under vegetation (e.g. flooding under forest canopies) is not always possible. Additionally, certain detection of partial surface water and/or inundated vegetation will be seasonally dependent as the vegetation cover changes. Thus, the field work aims to provide crucial data for validation and measuring performance baselines for DSWx detections.

At the highest level, the field work will be used to generate (a) high-resolution classifications in conjunction with very high-resolution optical imagery, mirroring those classes from the OPERA DSWx products² and (b) point data within the site extents that provide high temporal sampling of whether a point is inundated. Note that “partial surface water” and “inundated vegetation” will have overlap, thus the validation datasets will be generated to mirror the DSWx-HLS classes and the DSWx SAR suite. Such validation datasets will be acquired so that they are coincidental with relevant OPERA products and will be used to validate the appropriate requirements. This is described in detail in Section 4.5.

The field sites have been selected so that they address challenging retrieval areas, particularly for DSWx-HLS as discussed in [Jones, 2019] and [Jones, 2011], contain coastal areas including marshes that will be particularly relevant to validate the DSWx SAR suite. Further, the selected sites have either been previously visited or are easily accessible by United States Geological Survey (USGS) field teams. In other words, field experiments will be deployed in synchronicity with existing in situ data collection. The field sites and the site’s lead are summarized in Table 3.4. The field sites will be selected within the areas shown in Figure 3.2.

For the Validation of the DSWx product suite, the field team will deploy in situ sensors networks to survey a range of environments. All the field sites will yield wetted perimeters mapped by the field team twice annually where possible. Here, “wetted perimeters” refers to water classifications mirroring the DSWx product classes including areas with inundation under vegetation and partial surface water with

¹ Note that “partial surface water” detection performance is described in the requirements for DSWx-HLS (see requirement 3 for DSWx in 2.1) and “inundated vegetation” is described for DSWx-S1, DSWx-NI, DSWx-SWOT (see requirement 5 for DSWx in 2.1).

² See section 2.2.4 and Appendix B.

spatial resolution sufficient to test sub-pixel inundation declarations within DSWx products³. The field team will utilize GPS survey to precisely map these extents. Further, the field team will measure vegetation cover via ceptometer and/or densiometer. The former instrument estimates leaf area index and the latter canopy closure [[Garcia, et al. 2019](#)].

Additionally, temperature sensors will be deployed along transects and in contact with the soil substrate to provide high temporal point sampling of inundation. At a minimum, the temperature sensors will inform if a particular point is inundated during an input image acquisition date. When additional Validation image data are available and environmental conditions permit (e.g. cloud cover and vegetation density are low enough), these data can be combined to provide high-resolution classifications.

Depending on available resources and collaborator capacity, some sites may image field sites with Unmanned Aerial Systems (UAS) during the field team deployments. However, no validation strategies or activities rely on such data and thus the availability of such data will only enhance (not limit) Validation activities.

Although each field site will have its own set of challenges for field work, all the field data will be synthesized into the common formats specified above, namely as either (a) high-resolution classifications mirroring the classification from the OPERA DSWx product or (b) point data to determine whether a point is inundated. This data will be utilized in the validation activities that are described in section 4.5.

The DSWx Validation data will be published in peer-reviewed journals and made available to the public to benefit surface water monitoring efforts and scientific research including Validation activities for other NASA (e.g. NISAR and SWOT), regional studies at field sites, and leveraging OPERA field data for additional modeling and related detections.

Site Names	Region	Data Retrieved from Field Site	Field Team Lead(s)	Collaborating Agencies
Deschutes Wetland Complex, Sycan Marsh	Pacific Northwest	Wetted perimeters; Transects; Ground Photos; Airborne imagery (if available)	Adams (USGS Amphibian Research Initiative Program Lead) Jones (DSWx PI)	USGS USFWS TNC University of MT
Yellowstone National Park	Intermountain West	Wetted perimeters; Transects; Ground Photos	Jones (DSWx PI)	USGS NPS
Cottonwood Lakes	Northern Prairie Potholes	Wetted perimeters; Transects; Ground Photo	Bansal (Wetland Methane PI)	USGS
Leetown	Mid-Atlantic	Wetted perimeters; Transects; Ground Photos; Airborne imagery (if available)	Jones (DSWx PI)	USGS MD-DNR

³ This refers to requirement 1.

Patuxent National Wildlife Refuge	Mid-Atlantic	Transects, Ground Photos, Airborne imagery (if available)	Jones (DSWx)	USFWS

Table 3.4: List of DSWx field experiment sites, data collected, field team leadership, and collaborating agencies.

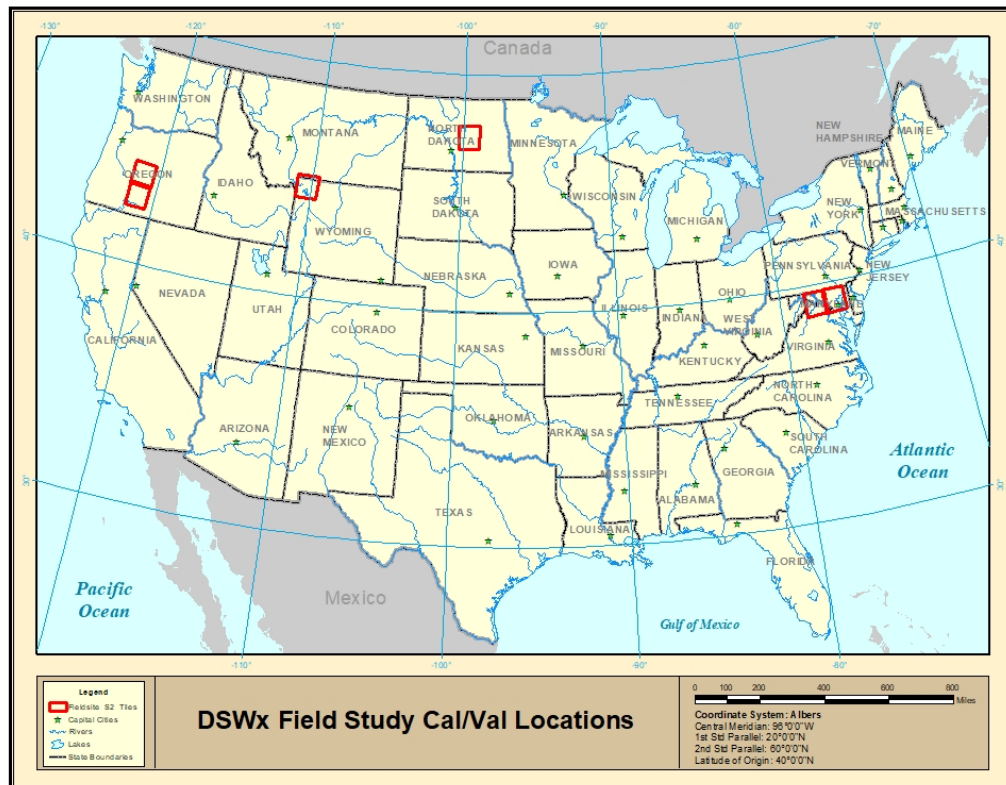


Figure 3.2: HLS tile boundaries covering field site regions proposed for Dynamic Surface Water Extent DSWx product Validation.

3.2.2.1 Deployment Timeline

The summary of the deployment timeline is in Table 3.4. The field work occurring *prior* to the DSWx-HLS validation period are at sites that will be used for partial surface water areas. That is, the field work relevant for DSWx-HLS will have been collected prior to its validation activities. Some of the validation activities occur after the validation period for DSWx-HLS (red cells in Table 3.4). Specifically, some “leaf-on” and spring campaigns occur during the DSWx-HLS validation period. This field work will provide expansive measurements of inundated vegetation so that sufficient data for DSWx-SAR validation will be available. However, there will be ongoing validation activities throughout the duration of OPERA and Validation data from fieldwork that will be integrated into validation activities across the suite of DSWx products even if certain products are not expected to be sensitive to certain classes.

Before deploying in situ sensors to all field sites, a pilot study will establish the standard operating procedures (SOPs) and field test the equipment that will be used at the other DSWx field experiment sites. The pilot study will be performed at the Mid-Atlantic site (Leetown Aquatic Ecology Science Facility) and will run from August 2021 through February 2022.

The field campaigns at all sites will occur twice annually for the following two years (FY 2022 and FY 2023). While all Validation data will inform DSWx development (including clarifying the performance baselines in the requirements described in Section 2.1), data collected through Validation field work will be directly comparable with DSWx-HLS and DSWx-S1 products as imagery is collected by those contributing satellite systems during the field campaign periods.

During FY 2022, equipment deployment and first surveys will take place in the spring, following melt⁴ and prior to leaf-out. FY 2022 second visits will occur late summer to early fall to allow use of the collected data in initial DSWx Validation. FY 2023 first site visits will occur post-melt, pre-leaf out to download data, check in situ sensor “health” and perform field surveys. Second visits in FY 2023 will be customized to capture conditions different from those encountered during FY 2022 field visits or by fall (before likely snow cover), whichever comes first. Whether in situ instruments are collected at the time of the FY 2023 second visit will depend on whether continued demand/support for fieldwork exists (e.g., for NISAR or SWOT mission support relevant to OPERA validation requirements) and/or collaborative agreements (e.g., with the TNC or Jones Center) are in place to allow longer-term operation. The DSWx Validation fieldwork deployment timeline is summarized in Table 3.5.

OPERA plans to cooperatively work with the NISAR and SWOT Validation teams for validation of OPERA DSWx product requirements, with details to be finalized as the project approaches those validation time periods.

⁴ Note that the requirements discuss water not ice and validation of icy, snowy areas is beyond the scope of the validation activities.

Activity	FY21	FY22				FY23				FY24				FY25			
DSW _x	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Fieldwork Pilot/Protocol development	x	x															
Pre-leaf out campaigns and equipment deployment			x														
Periodic image capture via commercial satellite	x	x	x	x	x	x	x	x	x								
Leaf-on, pre-DSW _x Validation I data collection				x													
Spring campaigns and equipment check							x										
Leaf-on Fall campaigns									x								

Table 3.5: Summary of DSW_x Field deployment. The red cells reference the validation period for DSW_x-HLS from Table 3.1. Light yellow cells refer to the validation period of DSW_x-S1 and DSW_x-NI; light blue, for SWOT. The continued field deployments after DSW_x-HLS will not affect the validation of the product requirements (see text).

3.2.3 Disturbance Field Work

The disturbance field work will be used to *assess* the *quality* of the DIST's current vegetation cover indicator layer with respect to in-situ vegetation cover. Although the field work for the DIST product will not be used to validate the DIST product requirements, the vegetation cover indicator layer is utilized by the disturbance detection algorithm. Therefore, its assessment is invaluable for algorithmic calibration (see the ATBD for details) as well as for users who may use this layer for correlative analysis. The assessment will give a rigorous statistical comparison of the vegetation indicator layer and the in-situ vegetation cover as determined by this field work. We will describe this assessment process in Section 4.6 and the related deliverables in Section 4.7.

The sites will be collected based on the expertise and experience of the UMD team that has built similar vegetation density products previously (including the MODIS VCF [[VCF User Guide](#)]). Some related field work and site selection done previously is discussed in the work of the [[Ying et al. 2017](#)]. There will

be at least 30 sites collected across three regions (10 different sites from each region) of the United States: Mid-Atlantic, Southwestern, and the Midwest. The Mid-Atlantic will focus on semi-urban areas; the Midwest, on agricultural land use; and the Southwest, on semi-arid vegetation ecotones. Field sites will be sampled seasonally, as feasible, to capture variations in intra-annual vegetative cover.

We now describe the precise data that will be generated from the field work. For each field site, the field team will select a 90 m square corresponding to a 3 by 3 grid of HLS pixels. Vegetation cover is defined in the requirements (Section 2.1) as well as Product Description Document [AD3] as the amount of plantlife orthogonal to the surface visible to the optical sensor. Sites will focus on sites with a mix of vegetative/non-vegetative covers.

The vegetation cover will be estimated at the sub-meter resolution available and then resampled to the data grid described above. The estimate will be determined using a variety of field techniques combined with high-resolution imagery described in [Ying et al. 2017]. If available, low altitude UAS multi-spectral data will be acquired over lands permitting UAS deployment by NASA.

3.2.3.1 Deployment Timeline

UMD will perform fieldwork in support of OPERA's validation of the disturbance product over the course of a year. This timeline is short because (a) the field work will be used to assess the intermediate vegetation cover estimate layers that are used in the disturbance product and (b) the field sites are selected based on previous UMD experience collecting related field data at the domestic field sites. The field data will not be used to encompass the multitude of scenarios required for verifying formal requirements although it will represent important thematic land uses described above.

UMD's work with respect to the field work will include a) *selecting* the field sites that provide a range of conditions and land cover types that are necessary for the near-real time global disturbance product, b) *collecting* the percent vegetation cover for 90 m by 90 m areas across at least 30 sites using multi-spectral cameras temporally coincident with the optical sensors Landsat and Sentinel-2, c) *developing* a format for the data to be compared with disturbance retrieval algorithms from optical/SAR imagery, and d) *characterizing* the vegetation cover conditions that impact the final disturbance models to quantify uncertainty in the final products.

Activity	FY21	FY22				FY23				FY24				FY25			
DIST	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Fieldwork Pilot/Protocol development	x																
Mid-Atlantic sites		x															
Midwest sites			x														
Southwest sites				x													

Table 3.6: Field work for DIST product. The red cells reference the validation period for DIST from Table 3.1.

3.3 Resource networks

Resource Network	Relevant OPERA product	Format	Coverage	Temporal Sampling	Description
<i>EDEN Gauge Network (3.3.1)</i>	DSWx	Pointwise	South Florida	Hourly	>150 gauges providing water surface heights and including contextual information on nearby topography and vegetation cover.
<i>USGS Gauge Data (3.3.2)</i>	DSWx	Pointwise	United States	Hourly	Real-time gauge data recording hourly hydrological variables include surface water height.
<i>UNR, UNAVCO, and JPL SGP GNSS data (3.3.4)</i>	DISP	Pointwise	Global	Sub-daily to daily	Processed and corrected 3-component displacement time-series from third party GNSS networks
<i>USGS GNSS data (3.3.4)</i>	DISP	Pointwise	United States	Daily	Processed and corrected 3-component displacement time-series from third party GNSS networks
<i>NISAR wetland Validation data (3.3.5)</i>	DSWx	Pointwise and Raster	N. and S. America and India	At time of NISAR overflights during Validation periods	Field measurements at NISAR wetlands Validation sites currently planned for summer 2024
<i>Existing Corner Reflector Arrays (3.3.3)</i>	CSLC	Pointwise	N.America and Australia	N/A	Accurately surveyed targets of known SAR brightness.

Table 3.7: Summary of Resource Networks and which OPERA products will utilize their data for validation.

3.3.1 Everglades Depth Estimation Network (EDEN)

The Comprehensive Everglades Restoration Plan (CERP) through the U.S. Army Corps of Engineers and the U.S. Geological Survey Greater Everglades Priority Ecosystem Science Program supports the Everglades Depth Estimation Network (EDEN) to provide consistent, documented, and readily accessible hydrologic and ground-elevation data for the Everglades (<https://sofia.usgs.gov/eden/>) for use in field analysis and habitat modeling [Jones et. al. 2012].

The EDEN includes an integrated network of more than 200 water-level gauges with approximately 160 broadcasting data in real-time [Haider, et al. 2020, USGS specification]. In combination with interpolation models and applications, the EDEN generates daily water-level data and derived hydrologic data across the freshwater part of the greater Everglades landscape. This data will be used for Validation data of the DSWx product. Specifically, it will provide near real-time point data on whether areas with mixtures of water and vegetation below DSWx resolution and/or areas under vegetation are inundated. This will be most relevant to the validation activities described in Section 4.5.

3.3.2 USGS river and tide gauge network

The USGS data will be used for DSWx validation activities.

For more than 125 years, the USGS has monitored flow in selected streams, rivers and coastlines across the United States and does so in cooperation with over 850 federal, state and local agencies. More information may be found at the [USGS National Water Information System webpage](https://maps.waterdata.usgs.gov/mapper/) (<https://maps.waterdata.usgs.gov/mapper/>)

The USGS Tidal Gauge Network accessed via the Water Quality Data Portal [Read, et al. 2020] provides pointwise data of numerous hydrological variables including water surface height, bathymetry, and estimated discharge.

The USGS tide gauge network will provide information about water level heights for DSWx validation activities. While the data will not be directly used for validation of OPERA products, the USGS data will inform whether other project-generated Validation data are in agreement with this real-time network of gauges.

3.3.3 Existing Corner Reflector Arrays

We are leveraging existing corner reflectors (CR) deployed in North America for missions such as NISAR and NASA's UAVSAR airborne SAR to assess the geolocation quality for both Sentinel-1 (C-band) and NISAR (L-band) co-registered SLCs (CSLCs) and for the RTC-S1 products. The Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV) has endorsed several test sites to perform calibration and validation activities. The SAR subgroup emphasizes its mission to "... foster high-quality synthetic aperture radar data from airborne and spaceborne systems through precision calibration in radiometry, phase and geometry, and validation of higher-level products." Table 3.7 summarizes the CR resource networks and their suitability for L- and C-band validation. Note that although the CSLC products have a North America geographical scope, the Australian Corner Reflector Array will be used primarily to validate Sentinel-1 CSLCs as this network is currently designed and recommended for C-band calibration/validation.

Site Name	Relevant OPERA Products	Region/ Country	Description	Owner
Rosamond Calibration Array	CSLC, RTC	California	20+ Trihedral Corner Reflectors Ka-band (0.7 m) L-band (2.4 m) and P-band (4.8 m). Supports Ascending observations. Designed for UAVSAR. Spans 10 km.	NASA/JPL
ASF Corner Reflector network		Alaska	~4 Trihedral Corner Reflectors > 2.8 m (2023) Intended for L-band but testing C and S-band RCS. Will support left looking observations. Designed for NISAR.	NASA/ASF

Oklahoma Corner Reflectors		Central US	10+ Trihedral Corner reflectors 2.8 (2023). Intended for L-band but testing C and S-band RCS. Should support right and left looking observations. Designed for NISAR.	NASA/Univ of Oklahoma
Australia Corner Reflectors		Queensland, Australia	40+ mostly C-band Trihedral Corner Reflectors for right looking observations. This array is regularly used for evaluating Sentinel-1 calibration. Until the evaluation of the NISAR reflectors is completed, this is the recommended calibration array for Sentinel-1.	Geoscience Australia

Table 3.8: Summary of Validation corner reflector networks.

3.3.4 UNR, UNAVCO, USGS and JPL SGP GNSS data

DISP validation will utilize dense continuous GNSS measurements in North America that are publicly available from repositories such as the [University of Nevada Reno](#) [Blewitt, et al., 2008], [USGS](#), [UNAVCO](#), and [JPL Space Geodesy Program \(SGP\)](#). These repositories have access to important L2 position data from third-party networks including EarthScope's Plate Boundary Observatory (PBO), Hawaiian Volcano Observatory (HVO), and the Continuously Operating Caribbean GPS Observational Network (COCONet). These networks are located in actively deforming regions of various sources (i.e. fault movement, volcanic deformation, hydrological changes, and etc.), such that performing the validation in these areas can be representative for the entire North America. We summarize them in Table 3.9.

GNSS Network	Relevant OPERA product	Coverage	Description
<i>PBO GNSS Network</i>	DISP	Mostly Western US	>1000 cGPS stations Near-real-time data accessible through UNR, UNAVCO, USGS, JPL SGP
<i>COCONet GNSS Network</i>		Caribbean	>85 cGPS stations funded by NSF; Network in tropical environment Near-real-time data accessible through UNR, UNAVCO, JPL SGP
<i>HVO GNSS Network</i>		Hawaii	58+ cGPS Near-real-time data accessible through UNR, UNAVCO, JPL SGP

Table 3.9: Summary of important GNSS networks over North America.

3.3.5 NISAR wetland Validation data

The NASA ISRO Synthetic Aperture Radar (NISAR) mission is planned for launch sometime in 2023. This data will be used for the DSWx product suite, specifically for the inundated vegetation and partial surface water validation activities. Its use for Validation will depend on the availability of this data during the validation period, and as such, OPERA will not necessitate this data for validation of the DSWx product. One of its science measurement requirements is for NISAR to be capable of classifying inundation extent in wetland areas (including both areas of open water, and surface water with vertically

emergent vegetation). This requirement will be validated by collecting in situ measurements and remote sensing imagery over a small number of wetland sites including:

Bonanza creek APEX site, Alaska, USA

Mississippi river and coastal marsh, Louisiana, USA

Everglades National Park, Florida, USA

Pacaya-Samiria, Peru

In conjunction with ISRO:

Chilika Lagoon, Odisha, India

Nalsarovar, Gujarat, India

Nanda Bet, Gujarat, India

3.4 Ancillary Data Products

There are a number of ancillary datasets that are used for product Validation data including high-resolution imagery and digital elevation models (DEMs). If the products are not public (e.g. commercial data where the license prevents sharing outside the Validation team), the datasets will not be published and only derived products will be available for public access. Below, Table 3.10 summarizes the ancillary products that will be used. Note forthcoming NASA Imagery and Validation data will be used depending on availability and will not impact deliverables or schedule. Determining such synergies is currently beyond the scope of the current version of the Product Validation Plan.

Ancillary Dataset	Relevant OPERA product	Format	Coverage	Temporal Sampling	Description	Availability	Spatial Resolution (m)
<i>Planet Labs PlanetScope and RapidEye</i>	DIST, DSWx	Raster Image	Global	1 Day	4 band (RGB+NIR) orthorectified for geospatial analysis of Planet Lab's Dove CubeSat Imagery. RapidEye will be retired in 2020 and PlanetScope will be the newer Planet Labs product.	Obtained via NASA Commercial Smallsat Data Acquisition (CSDA) program.	3.7 m
<i>Copernicus GLO-30</i>	RTC, DSWx	Raster Image	Global	N/A	Global Digital Elevation Model derived from a variety of sources including TerraSAR-X	Open dataset	30 m
<i>Height Above Nearest Drainage</i>	DSWx	Raster Image	Global	N/A	Maps height difference to map drainage based on DEM. Resolution based on available	Open datasets	10 m (North America) or

<i>(HAND)</i>					DEM. North America resolution relies on 3DEP and Global resolution on GLO-30		30 m (global)
<i>MAXAR Worldview</i>	DSW _x	Raster Image	Global	97 minute repeat pass	High resolution optical imagery from Maxar	Obtained via USGS administered Commercial Remote Sensing Space Policy	30 cm
<i>Sentinel-1 RTC via ASF</i>	RTC	Raster Image	Global	6 Days (Same as Sentinel-1)	Images are processed on-demand using the GAMMA-RTC workflow.	Freely available processing workflow	30 m
<i>NASA Imagery and Validation Data</i>	DSW _x , DIST	N/A	N/A	N/A	Leverage data from Validation and Monitoring activities related to Delta-X, SWOT, NISAR, AirSWOT, UAVSAR	Depends on availability and date of publication - will not delay OPERA	N/A

Table 3.10: Summary of ancillary products for Validation.

3.4.1 High-resolution Imagery from Commercial Satellites

High resolution imagery will be used as a primary source generating validation datasets for DSW_x and DIST. The two examples below are licensed and only derived analysis products will be published.

3.4.1.1 Planet Labs RapidEye and PlanetScope

The company Planet Labs has maintained a large constellation fleet of cubesats to rapidly image the globe since 2014 [[Boshuizen, et al. 2014](#)]. As of 2020, the Dove constellation has 120 active satellites that provide daily imagery of the Earth [[press release](#)]. The RapidEye product (retired since March 2020) and it's next generation release PlanetScope provide 3.7 meter resolution ortho-rectified analytics products that are processed to provide 4 band (RGB + NIR) imagery that can be used for geospatial analysis [[specification](#)]. Planet data will be accessed through the existing NASA Commercial Smallsat Data Acquisition (CSDA) program.

3.4.1.2 MAXAR

Submeter panchromatic and slightly coarser resolution multispectral imagery (VIS-NIR-SWIR) from MAXAR [Optical Imagery \(maxar.com\)](#) will be requested and acquired through USGS administered Commercial Remote Sensing Space Policy ([Commercial Remote Sensing Space Policy \(usgs.gov\)](#)) via the NextView and NorthernView licensing mechanisms. Specifically, the 30 cm resolution [Worldview-3](#) products will be utilized. These can be obtained from the existing MAXAR archive or through special tasking via Federal licensing managed by USGS/NGA.

3.4.2 Copernicus GLO-30 Digital Elevation Model (DEM)

The Copernicus GLO-30 DEM is a global DEM that will be used for validation activities related to the CSLC, RTC, and DSWx.

The European Space Agency's earth observing program Copernicus created this global 30 meter DEM and has made it publicly available including via an open S3 bucket. The GLO-30 DEM is derived primarily from TanDEM-X data with void data being filled in from a variety of sources including SRTM [[Product Handbook](#) - <https://doi.org/10.5069/G9028PQB> Accessed: 2021-07-29].

3.4.3 Height Above Nearest Drainage

Height Above Nearest Drainage (HAND) will be used in Validation activities related to DSWx SAR-based products. More specifically, it will be used to help enhance the final validation datasets that will be compared to OPERA products to ensure that inundation extents agree with elevations (i.e. those from the GLO-30 DEM described in the previous section) at the validation site.

HAND is a topographic and hydrologic index in which each pixel stores the height difference between the pixel and its drainage. It can also be considered an elevation model normalized by the height of the drainage network inside a basin [Nobre et al. 2011, Rahmati et al., 2018]. Several HAND datasets are available to the public. Deltares released a global 30 m HAND layer in 2016 on [Google Earth Engine](#) [Donchyts et al. 2016]. Researchers at Oak Ridge National Laboratory developed a [10 m HAND layer](#) for continental US [Liu et al., 2020] based on the USGS 3DEP DEM [Arundel, et al. 2016]. A new global HAND layer has also been developed at the Alaska Satellite Facility, by processing the Copernicus GLO-30 DEM [Leister-Taylor et al., 2020], and HydroBasins using [PySheds](#) [Lehner and Grill, 2013; Bartos, 2020].

3.4.4 Alaska Satellite Facility (ASF) GAMMA Sentinel-1 RTC products

The ASF Gamma Sentinel-1 RTC products will be used to validate OPERA RTC-S1. The GAMMA software [Werner et al. 2000] is a (licensed) SAR processing software used for a variety of SAR processing. OPERA will use the open source software package ISCE3 [Rosen et al. 2018]. There may be small differences in the implementation of the RTC processing and these will be discussed in Section 4.3. However, the RTC imagery will provide an important basis for comparison for the OPERA products as GAMMA has been a popular software tool for Radiometric Terrain Correction of SAR imagery.

The GAMMA-corrected RTC imagery of Sentinel-1 data will be obtained from the NASA Alaska Satellite Facility (ASF) DAAC. ASF provides on-demand, analysis-ready Sentinel-1 data including RTC imagery through its HyP3 service [documentation]. For RTC products, a user selects Sentinel-1 SLC data, which triggers a cloud-based processing chain that generates the ASF RTC image. The specific workflow is documented in the ASF on-demand [user guide](#) and in an accompanying [ATBD](#) document. The RTC Products are provided at 30-m resolution that matches the GLO-30 DEM used for georeferencing.

3.4.6 Global Forest Change 2000–2019

Landsat-derived global tree cover and forest disturbance maps from [Hansen, et al. 2013] will be used to determine forested areas for RTC validation activities. Specifically, the tree cover layer will be used to map forest areas (see the requirements in Section 2.1). Tree cover in this product is defined as the percent of a 30 meter Landsat pixel containing canopy.

3.4.7 NASA Imagery and Validation Data

NASA has committed to an open data policy [\[link\]](#) and has a number of initiatives to make its data products more accessible including [\[Plato, 2019\]](#), [\[Smith, et al. 2020\]](#), and [\[McGibbney, et al. 2016\]](#). Again, when available, we will utilize available Validation data from UAVSAR, AirSWOT, DELTA-X [\[Simard, et al. 2020\]](#), SWOT, and NISAR to complement our existing calval activities.

While there are synergies between ongoing NASA missions and the OPERA project (particularly DSWx and DIST), the aggressive schedule of OPERA may not permit validation activities to utilize such data as it may not be publicly available in time for the OPERA validation activities. While there are ongoing discussions between relevant external Validation teams, additional NASA Validation data and NASA imagery will only enhance (not limit) validation activities for OPERA when it is available and is currently beyond the scope of the current validation activities. We will document and record relevant updates of Validation activities in future versions of this document.

3.5 OPERA Validation Science Data System (SDS)

The OPERA Validation SDS broadly refers to the computing infrastructure that supports Validation activities including on-premises resources and a validation database. The OPERA Validation SDS will (a) collect up-to-date validation data across resource networks and other ancillary data sources, (b) store the relevant validation data discussed in this section in a database, (c) store provisional OPERA products used in validation (Section 2.3), (d) instantiate validation analyses via reproducible science code that will be developed in the public domain (as is the OPERA project policy), and (e) expose aforementioned data to the Validation team via an API. All the validation data housed in the OPERA Validation SDS and the final validated OPERA products will be delivered to a Distributed Active Archive Center (DAAC) for community benefit (see 3.5.1 regarding the OPERA Validation Database for more details).

In addition, the Validation SDS will include a clone of the production system that can produce provisional OPERA products for validating OPERA project requirements. This clone utilizes the final OPERA production algorithms, input datasets, and any additional ancillary datasets required for processing, while running on a system deployed precisely as the final OPERA Product SDS (see 2.3). This clone will generate products exactly as will be nominal for the project. This clone will be operated by the SDS team and the validation team will request products to the clone's operators. Once the products are generated in this clone of the Product SDS, the validation team will be able to access the final product using various APIs associated with this clone. Specifically, these APIs will serve two essential purposes: (a) to search through the generated provisional products using spatial and product metadata and (b) to obtain urls (and object store locations) to retrieve these provisional OPERA products. After the validation period, the SDS team will re-generate all the products used for validation activities using the final Product SDS and deliver these products to the DAAC.

The independent validation data (discussed earlier in Section 3.2 - 3.4) will be collected from the external providers and data archives (such as one of NASA's DAACs), then cataloged in the validation database (see the next subsection regarding details on the Validation database). The OPERA Validation SDS where the actual comparative analysis for verifying requirements is developed, deployed, and operated in the AWS Cloud, specifically US-West 2, Oregon region, where the DAACs host products from other missions that are inputs to the OPERA SDS. The OPERA Validation SDS will be managed by the Validation team. The component of the Validation SDS consisting of a mirror of the OPERA Product SDS will also be in the AWS cloud (US-West-2). This mirror of the Product SDS will be maintained and operated by the SDS team, while funded by the PST team.

3.5.1 The Validation Database

The OPERA Validation database will store and organize independent datasets used for validation, streamlining the validation activities and information sharing among members of the PST. The validation database provides a unified interface to:

- query values and locations of geodetic, hydrological, and biophysical observations while enforcing data access permissions to the PST and external collaborators.
- automatically updating ancillary data from remote repositories
- integrate land cover products to identify gaps in validation sites, locate outliers for further investigation, and plan data collections.

The validation data will be accessible using python APIs and can be analyzed for the verification of OPERA requirements. Each independent validation data product will be associated with attributes such as geophysical metadata, retrieval metadata parameters (including retrieval times), spatial extent, and a url to that dataset location.

We label all validation datasets as either source, intermediate, or validation products. Source datasets include in-situ data from discipline-specific networks (e.g. GNSS data) and high-resolution remote sensing (e.g. Planet imagery) that may be fetched from remote servers. Intermediate datasets are source datasets converted into GIS format and conditioned to facilitate the calculation of validation products. Validation products are intermediate datasets optimized for comparison with the provisional OPERA products generated by the OPERA Validation SDS. Intermediate and validation products are stored in separate tables, such that calculated products are kept separately from measured products in the database.

The validation database will work in conjunction with ancillary routines to dynamically fetch data from remote servers as orchestrated by the OPERA Validation SDS, reformatting or processing datasets to generate relevant intermediate or validation products.

Access to (or a clone of) the validation database will be provided to NASA DAACs. Additionally, validation products and validation datasets will be delivered to identified NASA DAACs for archival purposes according to requested schedules. While some source products such as Planetscope will not be archived due to copyright, identifiers for those products and derived validation products will be provided. Note that any intermediate data that cannot be used to re-create the original licensed source data can be published, e.g. a water mask obtained from the licensed source data.

3.6 Roles and Responsibility for Validation

Validation is a joint effort involving project team members. OPERA has well-defined product divisions with specialists committed to the validation of these products and the verification of algorithms that produce those products. Table 3.10 summarizes the activities and the roles and responsibilities of each team.

Description of the activity	PST**	ADT**	SDS**
Validation SDS			
Re-deployment of the OPERA Product SDS (i.e. the clone of Product SDS) to request the generation of OPERA provisional products utilizing up-to-date OPERA algorithm implementations, identical input datasets, and identical ancillary processing data.			x*

API access to search product metadata and retrieve download urls generated by the clone of Product SDS in Validation SDS in previous row			x
Development and maintenance of the Calval database to house independent validation data	x		
Verification of product algorithms	x	x	
Validation Activities			
Validation fieldwork	x		
Generation of validation datasets	x		
Production of project products for validation	x		x*
Comparison of results to requirements	x	x	
Re-generation of project products used in validation with delivery to the DAAC			x
Post-validation period			
Ongoing production of project products after validation period complete			x
Periodic validation assessment of the project products	x	x	

Table 3.12 Summary of Validation roles and responsibilities. *Indicates funded by the PST but where the workforce of SDS is leveraged for executing. **The UMD team is responsible for these tasks pertaining to the DIST product suite. Shaded text indicates those activities refined or added since V1.1 of this plan.

The OPERA Project Science Team (PST) is responsible for the overall Validation activities in this document such as generation of validation datasets, generation of the provisional products, and comparison of results to requirements. The University of Maryland (UMD) team will be responsible for these (including related cloud cost) as they relate to the DIST product. All DIST validation data will be backed up in a shared cloud Validation Database (see Section 3.5.1) so that this data can be used for other validation activities.

The PST will also work with the ADT on algorithm verification (where needed) where the specification of the algorithmic parameters are dependent on validation data sets. The PST will also plan the fieldwork and organize any logistics involved in deploying instruments or conducting field campaigns. They will also oversee the coordination of relevant activities amongst other teams. In the post-validation period, the PST will perform validation assessment of the project products. The specifics of the ongoing validation activities during this post-validation period is detailed in Section 4.7.

The Algorithm Development Team (ADT) will be in charge of the algorithm development, verification and calibration of algorithms, testing, and will work closely with PST on the assessment of the project products.

The Science Data System (SDS) team will be managing and operating the re-deployment of the Product SDS within the Validation Environment for the validation to request provisional OPERA products. The SDS team will make sure that the re-deployment can be searched, and products retrieved; as can be done in the Product SDS. The PST team is responsible for funding this deployment, the processing costs that are incurred, and the SDS workforce. The continuous production of validated output products after the Validation period is part of the SDS scope.

4 Validation Strategies

We now have all the material to thoroughly discuss the validation strategy of each OPERA product. The described validation data will be sufficient to validate respective requirements [AD2] introduced in Section 2.1. Specifically, for each OPERA data product, a Validation strategy is detailed comparing OPERA data products to relevant in-situ and ancillary datasets. The strategy includes the quantitative analysis that will be used to demonstrate the requirement is satisfied.

We remark that datasets including the OPERA data products, the auxiliary datasets, if in the public domain, and any intermediate products will be made publicly available for further algorithmic testing and validation.

The structure of this section is as follows. Any additional preprocessing of OPERA products or Validation data prior to validation is described in the next subsection (4.1). In some cases, such preprocessing or selection of OPERA datasets or Validation data will be performed to remove errors that are outside the scope of the OPERA project, e.g. quality flags in input datasets. The product strategies are discussed in Section 4.2 to 4.6. In Section 4.7, we summarize the concrete product and analysis deliverables from each of the validation activities that have been scoped out. In section 4.8, we describe the strategy for ongoing validation of the OPERA products.

Products	Requirements Verified (see section 2.1)	Validation Sites	Comments
<i>Coregistered SLC</i> (<i>CSLC-S1 and CSLC-NI</i>)	Absolute location error (ALE) assessment	Sites selected to obtain coverage over corner reflector arrays: <i>Existing Corner Reflector Arrays for ALE :</i> • Rosamond Corner Reflector Arra (California, USA) • ASF Corner Reflectors (Alaska, USA) • Oklahoma Corner Reflectors (Oklahoma, USA) <i>New OPERA Corner Reflectors:</i> • OPERA Corner Reflectors (Central Valley California, USA)	The 7-8 OPERA Corner Reflectors to be deployed over Central Valley California, USA.
	Relative location error assessment	Validation will occur over areas that contain both persistent scatters (PS) and distributed scatters (DS) with maintained coherence (i.e. temporal coherence > 0.5). The North America Sentinel-1 12 day coherence study (Figure 4.10, [Kellndorfer et al. 2022]) will be used as a reference guide to ensure that the selected validation sites meet the temporal coherence requirement.	In addition, validation activities will occur on overlapping areas covered by Sentinel-1 and NISAR so that users can directly compare CSLC-NI with CSLC-S1 products.

		<i>Validation sites for the RLE:</i> • Big Island, Hawaii, USA • Aleutians, Alaska, USA • Oklahoma, USA • Oregon, USA • California, USA • Virginia, USA • Texas, USA • Puerto Rico, USA	
<i>Radiometric Terrain Corrections</i> (<i>RTC-S1</i>)	Absolute and relative geolocation ensuring one site with high topography (1 & 3)	Sites with corner reflector arrays used for absolute geolocation accuracy: • Fairbanks, USA • California, USA • Oklahoma, USA Sites for relative geolocation accuracy: • Mato Grosso, Brazil • Arkansas, USA	Representative of broad land cover types (see Table 4.2).
	Flattening - ensuring one site has high surface slopes (2 & 3)	• Mato Grosso, Brazil • New Mexico, USA • Appalachian Mountains, USA	Representative of broad land cover types (see Table 4.2).
<i>Displacement</i> (<i>DISP-S1 and DISP-NI</i>)	LOS displacement rate (1 & 2)	<i>Validation Approach 1 (VA1):</i> Using GNSS Resource Networks (ensuring density of GPS stations is high enough for validation): • California, USA • Oklahoma, USA • Big Island, USA • Houston/Galveston, USA • Unimak Island, USA <i>Validation Approach 2 (VA2):</i> Same validation sites as in VA1 and additional negligibly deforming sites within North America are TBD. For the latter, validation will occur over negligibly deforming areas that contain both Persistent Scatterers (PS) and Distributed Scatterers (DS) in a scene with maintained coherence through time (i.e. temporal coherence > 0.5). The North America Sentinel-1 12 day coherence study previously done for NISAR (Figure 4.10) will be used as a reference guide to ensure that the selected negligibly deforming validation sites meet the temporal coherence requirement. <i>Validation Approach (VA3):</i> New OPERA Corner Reflectors: • OPERA Corner Reflectors (Northern	Using GNSS datasets from: EarthScope's Plate Boundary Observatory (PBO), Hawaiian Volcano Observatory (HVO), and Continuously Operating Caribbean GPS Observational Network (COCOONet).

		California, USA)	
<i>Surface Water Extent</i>	Open water detection (2 & 4)	Sites to be selected using stratified sampling of 20 km squares using up-to-date DSWE-HLS model to estimate current global water distribution as in [Pickens, 2020] .	Expected to obtain 30 5-km square chips of high-resolution data that is coincident to the HLS tiles
	Partially surface water and inundation under vegetation (3 & 5)	<i>OPERA in situ measurements:</i> • Deschutes Wetland Complex, Sycan Marsh, USA • Yellowstone National Park, USA • Cottonwood Lakes, USA • Cape Cod National Seashore, USA • Leetown Aquatic Ecology Center, USA • Patuxent National Wildlife Refuge	Sites of long-term wetland study (USGS, TNC) Wide variety of previously studied wetlands (USGS/NPS) In situ monitored prairie potholes (USGS) Previous marsh dieback research site (USGS) Controlled experiment capabilities (USGS) Site of long-term wetland study (USFWS/USGS)
<i>Disturbance</i>	Vegetation Cover loss (1)	Sites to be selected using stratified sampling of the 30m HLS pixel grid. Stratification is created from prototype disturbance product.	
	Current vegetation cover indicator (NA)	Sites to be selected by UMD in three regions: the Mid-Atlantic with a focus on semi-urban areas; the Midwest, on agricultural land use; and the Southwest, on semi-arid vegetation ecotones.	Expected to obtain 30 90m x 90m chips of high-resolution data that is coincident to the HLS tiles

Table 4.1: Summary of Validation validation sites that will be used for the various OPERA data products. These will be revisited again for each validation strategy in Section 4. During the discussion of validation strategies, we will discuss how sites were selected to help validate requirements.

4.1 Preprocessing for Validation

Some input datasets may contain errors whose calibration and correction are beyond the scope of the OPERA project. In the event that such errors within any datasets used for validation, additional corrections and ad hoc updates/selections will be performed. If after such preprocessing, the input dataset is still degraded, then efforts will be made to utilize different data to complete the necessary validation activities specified by the L2 requirement.

4.1.1 Clouds and Cloud Shadows

Currently, HLS input data include cloud masks [\[Claverie, 2018\]](#), using the LaSRC processor for Landsat-8 and the FMask for Sentinel-2. The LaSRC processor obtains 4.7% omission error (or 95.3% recall) and 24% commission error (or 76% precision) [\[Foga, et al. 2017\]](#). While the FMask data has not been completely validated, studies indicate a median of approximately 90% precision and recall with large variability across scenes (30% drop in some validation scenes) [\[Baetens, et al. 2019\]](#). As noted in [\[Claverie, 2018\]](#), cloud classification is more challenging in Sentinel-2 without the thermal band present in Landsat-8 data.

If Validation activities related to HLS input datasets require additional cloud masks or scene-wise heuristics due to cloud coverage, these additional cloud masks will be carefully recorded and documented.

During previous USGS DSWE algorithm research and development [Jones 2019], the aforementioned tendency toward commission errors in HLS cloud masking led to DSWE product modifications that allow the user to ignore HLS cloud masks for their application, and, if possible, employ more accurate cloud masking techniques. A similar approach will be taken during DSWx Validation.

For the DIST product suite, to mitigate cloud obfuscation, more advanced mask techniques and additional quality heuristics from [Potapov et al. 2020] will be applied both in production and for Validation activities, all of which will be carefully recorded and distributed.

4.1.2 Radio Frequency Interference (RFI)

Radio frequency interference (RFI) is a well-known problem for communications and imaging systems occurring in the microwave spectrum that degrades a SAR system. RFI degrades a SAR system when microwave signals operating at an overlapping frequency of the SAR system mix and produce incoherent returns [Shimada et al., 2009]. This interference can be the result of ground communication systems, television networks, air-traffic surveillance, and/or military-based radiation. RFI can often be readily seen as bright linear strips parallel to the range direction (Figure 4.1) in the amplitude image of SLC data or bright linear features in the azimuth-time range frequency diagram of a SAR dataset. RFI lowers the signal-to-noise ratio.

Sentinel-1 images can be severely affected by C-band RFI [Monti-Guarnieri et al., 2017] (Figure 4.1). Only recently, Sentinel-1 RFI mitigation is performed utilizing auxiliary data from the Sentinel-1 Instrument Processing Facility (IPF) [Piantanida et al., 2016 (updated 2021)]. The NISAR processing mitigates RFI based on a time-frequency analysis of the raw SAR data and SLC image. Still, due to the diversity of RFI signatures, it is likely that some RFI will remain in the final processed NISAR data.

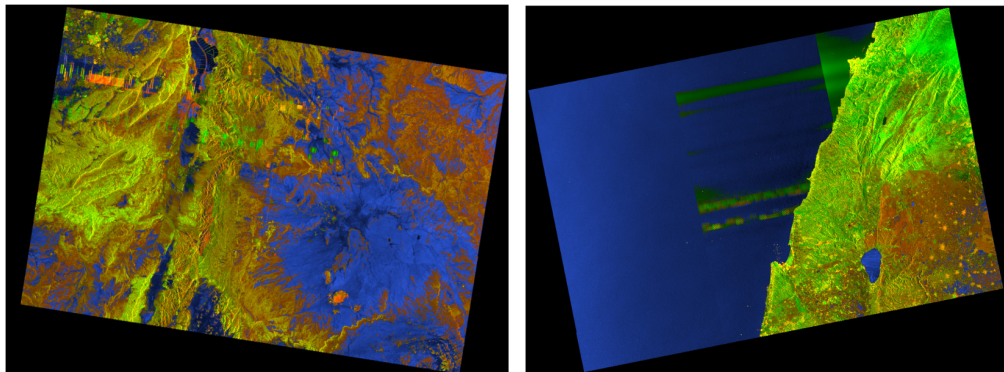


Figure 4.1: RFI in Sentinel-1 data over the Middle East. Copernicus Sentinel data 2021, processed by ESA.

Typically, incoherent RFI can be reduced with a matched filter applied in the focused SAR image due to the large coherent signal-processing gain of the matched filter operation. Despite this mitigating effect, high-power RFI may remain in the focused image as spatially correlated signal noise. Indeed, uncompensated RFI reduces the SAR image quality, affects the integrity of polarimetric signatures, and increases the noise level of the SAR-phase observations [Rosen et al., 2008; Meyer et al., 2013].

The OPERA project will not perform additional RFI corrections as it assumes the input data to be validated and calibrated by ESA and NASA for Sentinel-1 and NISAR projects respectively. If RFI

severely degrades an input dataset for Validation, efforts will be made to utilize different input imagery within the input dataset catalog.

4.1.3 Propagation delays

Synthetic Aperture radar imagery is impacted by the ionosphere and troposphere due to phase delays that occur in those portions of the atmosphere [Hanssen 2001]. In the ionosphere, the electromagnetic signal interacts with free electrons of the ionosphere; in the troposphere, the signal is slowed due to variations in pressure, temperature and relative humidity. The ionosphere is a dispersive medium, hence ionospheric propagation delays in L-band are approximately 4 times larger than for C-band. The ionospheric delays can be estimated directly from the SAR data [Fattahi et al. 2017]. A correction layer is included in the DISP-NI product, while this is still TBD for the DISP-S1 product. For the tropospheric delay, the OPERA project will leverage weather model data to estimate pressure, temperature, and relative humidity over a SAR acquisition and then correct phase using physical models with the weather model data as input using RAiDER processing tools [Bekaert et al. 2021]. The recent [comprehensive study](#) done as part of the NISAR Tropospheric Noise Working Group identified the greatest reduction from tropospheric variability from the ECMWF HRES model both at the global and continental scale. The study compared the ECMWF ERA5 reanalysis product, the NCUM analysis model of National Centre for Medium Range Weather Forecasting, and the GEOS analysis model by the GSFC Global Modelling and Assimilation Office (GMAO). The HRES model is a commercial model provided at 9 km with 5 hour and 45 minute temporal resolution. Negotiations by the NISAR project have already begun to determine the feasibility of licensing the use of this commercial product; if not feasible for OPERA, another model from this list will be selected. Such corrections may be applied if not done operationally during DISP Validation activities to remove phase propagation errors.

4.1.4 General Image Anomalies

While systematic OPERA Quality Assurance (see section 1.2) will exclude a majority of image anomalies, image and/or metadata anomalies occasionally will avoid detection and correction (e.g. data dropout, corrupted metadata, or spurious banding), and could impact the quality of OPERA Validation activities. Therefore, specifically for validation of products at OPERA Validation sites where image anomalies were not excluded, the mitigation strategy will be to identify any such anomalies as they occur in OPERA products and validation datasets. Additional heuristic and ad hoc corrections will be made for Validation when image anomalies are identified (or affected input datasets will be filtered). Such anomalies and associated corrections will be well documented and shared.

4.2 Co-registered SLC product

We are performing two sets of analyses to validate the CSLC requirements from Section 2.1:

1. *Relative location error (RLE)* assessment (for CSLC-S1 and CSLC-NI, respectively)
2. *Absolute location error (ALE)* assessment (for CSLC-S1 and CSLC-NI, respectively).

The validation strategy for these assessments will be identical for the S1 and NI OPERA products. Note that for both approaches, we will validate each requirement separately for the ascending and descending tracks from S1 and NI.

The *relative location error* (RLE) can be evaluated from each stack of ascending and descending CSLCs following state-of-the-art methods in geolocation accuracy assessment for stack coregistration (e.g. [Yunjun et al., 2022], [Fattahi et al., 2016], [Lei et al., 2021]). The accuracy of determining the misregistration between two SAR images is known to decrease as temporal decorrelation increases due to the increased time difference between the pairs. The issue can be circumvented by using a fully connected network of small temporal baseline image pairs that overlap in time to obtain an unbiased time series of misregistration/offset [Fattahi et al., 2016]. For each of the CSLC stack (spanning the start of product record to recent acquisition, see Table 2.1), we will pair each CSLC image with its N -neighbors in time as illustrated in Figure 4.2a (number of neighbors, N , TBD). Then, we will compute dense grids of sub-pixel differential offsets by cross-correlating each of the image pairs and apply a simple least squares method to obtain the time series of RLEs between each reference image and the secondary image within the stack. The resulting RLEs in the range and azimuth directions for the ascending and descending tracks will then be statistically analyzed (e.g. L2-norm residual analysis). Unreliable pixels (i.e. pixels that have large residuals or have low coherence values over time) are discarded based on a predefined threshold value (TBD) prior to comparing against the requirements. This comparison will provide metric offsets in each coordinate direction to determine if the requirement is satisfied.

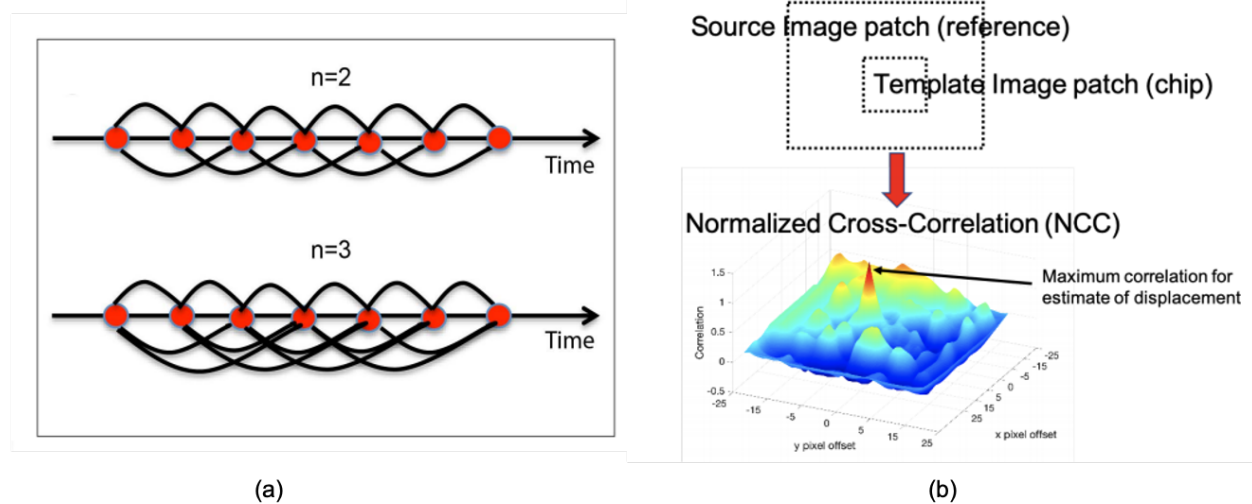


Figure 4.2: (a) Examples of the sequential network of CSLCs, with circles showing the SAR acquisitions and arcs representing the pairs of CSLCs [Fattahi et al., 2016]. (b) Temporal matching using normalized cross-correlation, from [Lei et al., 2021].

To evaluate the *absolute location error* (ALE) of the CSLC product, we will use corner reflector (CR) arrays (such as the OPERA CRs shown in figure 3.1) listed in Table 4.1. We will evaluate the ALE for each CSLC image in each of the stack to obtain a time series of offset. For each of the CSLC, the image coordinates of the CRs are identified at sub-pixel accuracy using point-target analysis and then compared to field-measured CR coordinates (i.e. [Gisinger et al. 2021], [Schubert et al. 2017]). For example, Figure 4.3 shows the estimated amplitude peak of the CR arrays for Sentinel-1 A/B for a corner reflector in California; this can be compared to the geolocation of the specific corner reflector and determine if the relevant requirement is met.

For the CRs discussed in Section 3.2.1, geolocation information is obtained from the co-located cGNSS stations. We will use standard GPS and laser surveying techniques for measuring the position of the vertex

of the reflector. The average of the measurements will then be directly compared to the point-location in the CSLC image to determine any offsets. For the other CR arrays, position data can be accessed through CEOS or on their public repositories (e.g. [Rosamond CR array](#)).

The CSLC validation sites are listed in Table 4.1 and are displayed graphically in Figure 4.4. RLE validation sites are represented by polygons whereas ALE validation sites are marked by stars. The CSLC validation sites are selected such that they represent the North America geographic scope while leveraging existing corner reflector arrays to evaluate the geolocation requirements.

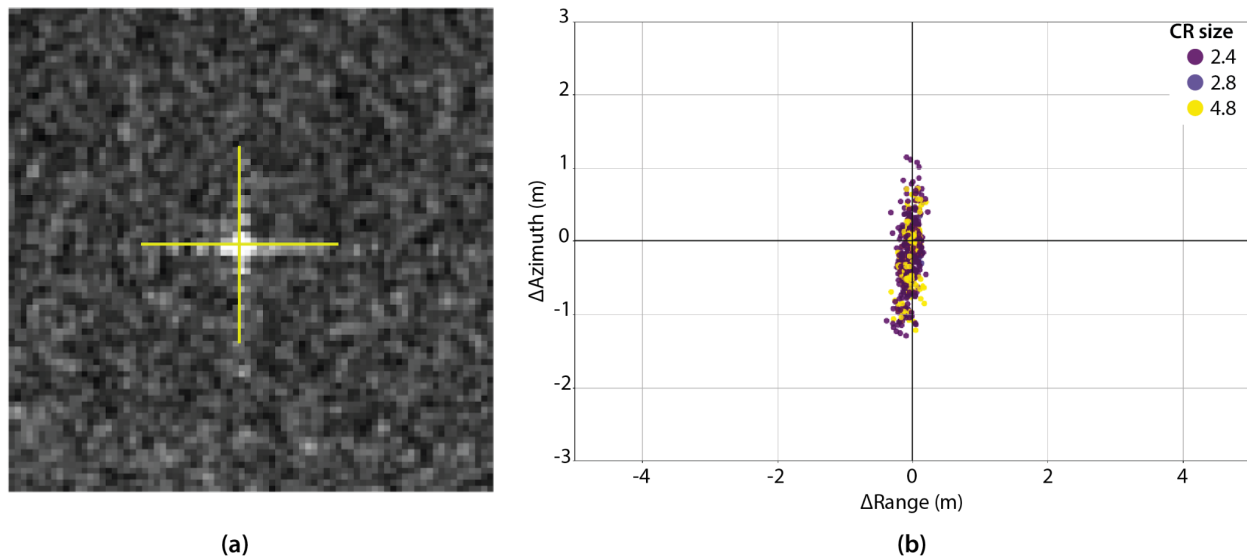


Figure 4.3: (a) One of the 2.4-m corner reflectors in Rosamond, CA, as seen from a Sentinel-1 descending track CSLC with the estimated phase center (yellow cross). (b) Example of an absolute location error analysis for the Rosamond CR array using a stack of Sentinel-1A/B CSLCs spanning Jan 6, 2021 to Dec 14, 2021.

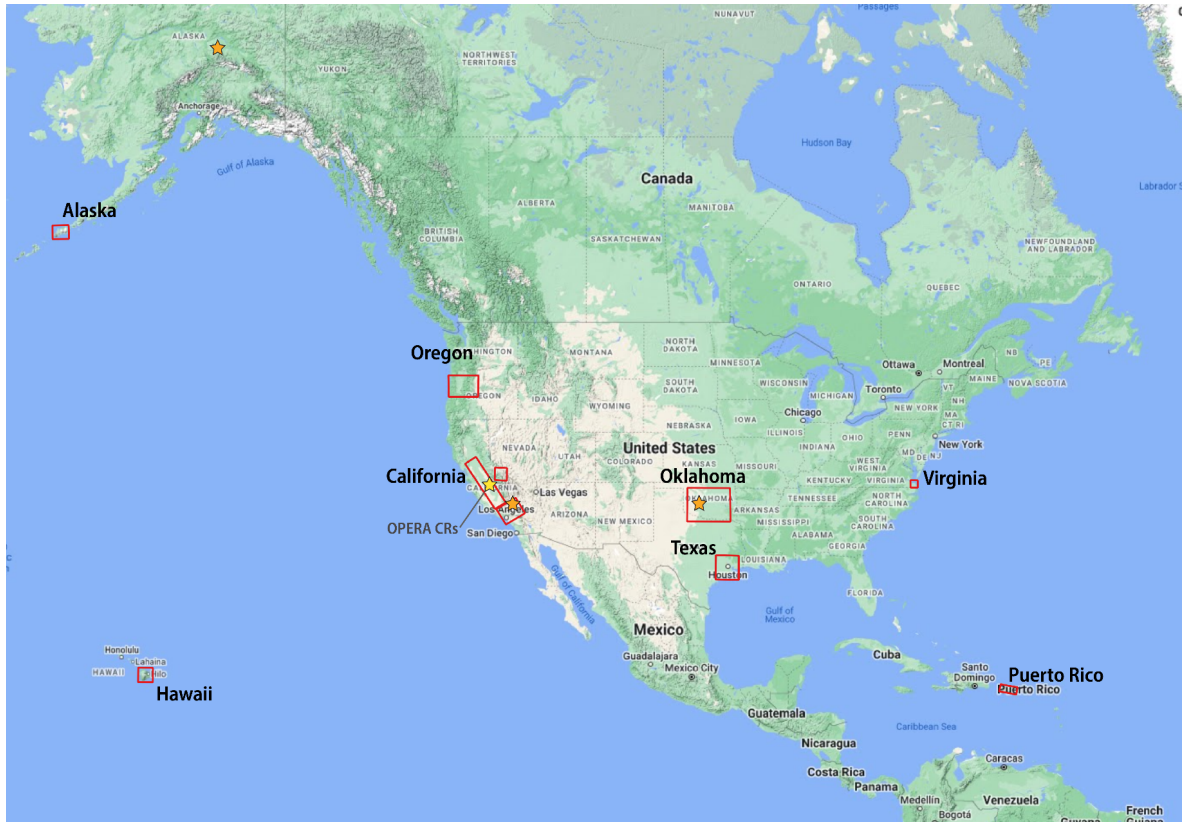


Figure 4.4: CSCLC validation sites over North America as listed in Table 4.1. Red polygons are validation sites for RLE whereas the orange stars are validation sites for ALE. The proposed location of the CRs to be deployed by OPERA in Central Valley, CA is marked by a yellow star. The CSCLC validation sites are chosen to represent diversity of targets in North America and where existing corner reflector networks can be leveraged.

4.3 Radiometric Terrain Correction Product

We are performing two different analyses to validate RTC-S1 products to ensure they meet the requirements specified in Section 2.1. These two analysis tasks are (1) validation of geolocation accuracy (both relative and absolute) and (2) validation of the radiometric terrain flattening.

First, we discuss task 1, the validation of the requirement related to geolocation: both *absolute geolocation* and *relative geolocation*. Accurate geolocation of RTC-S1 products is essential to ensure alignment of SAR products with data from other sensors and to allow for the derivation of unbiased parameters from time series of RTC-S1 products such as statistical parameters or time series information. The geolocation requirement for RTC-S1 products is specified in Section 2.1 of this document. Note that the requirement will be less stringent than that specified for the CSCLC product as CSCLC requirements are driven by interferometry which requires a tenth of a pixel or better relative alignment. That said, the geolocation validation strategies utilize a similar methodology as those described in the previous section for validation strategies of the CSCLC requirements.

To evaluate the *absolute geolocation accuracy*, as in the case of CSCLC, we consider ascending and descending stacks of RTC-S1 stacks over corner reflector sites (see sites from Table 4.2 labeled “Geolocation Accuracy” and are mapped with the other validation sites in Figure 4.6). For these sites, we

evaluate geocoding accuracy at the sub-pixel level using a set of corner reflectors distributed across the image frame. From these measurements, statistical metrics for geolocation accuracy are derived for data in ascending and descending orbit direction, and compared against requirements.

For the *relative geolocation accuracy*, we will analyze the alignment of time-series of RTC-S1 using standard subpixel co-registration techniques as in [Lei et al., 2021] or [Guizar et al. 2008]. This will be done at the validation sites labeled “Coregistration” in Table 4.2 and mapped in Figure 4.6. Specifically, each image in a stack will be compared to a single reference image and subpixel offsets will be determined from each comparison. These offsets will be used then to verify the *relative geolocation accuracy* is satisfied.

Finally, we discuss the validation strategy for the *radiometric terrain flattening quality*, i.e. task 2. The radar-cross section (RCS) of a distributed target varies with the geometric area of the illuminated pixel (described by Small [2011]), and also with incidence angle (depending on the landscape characteristics). To verify whether the radiometric terrain flattening process has indeed “flattened” the image backscatter, we outline two assessments proposed in Small [2011] and Truckenbrodt et al. [2019], which are assessments 1 and 2 below, respectively.

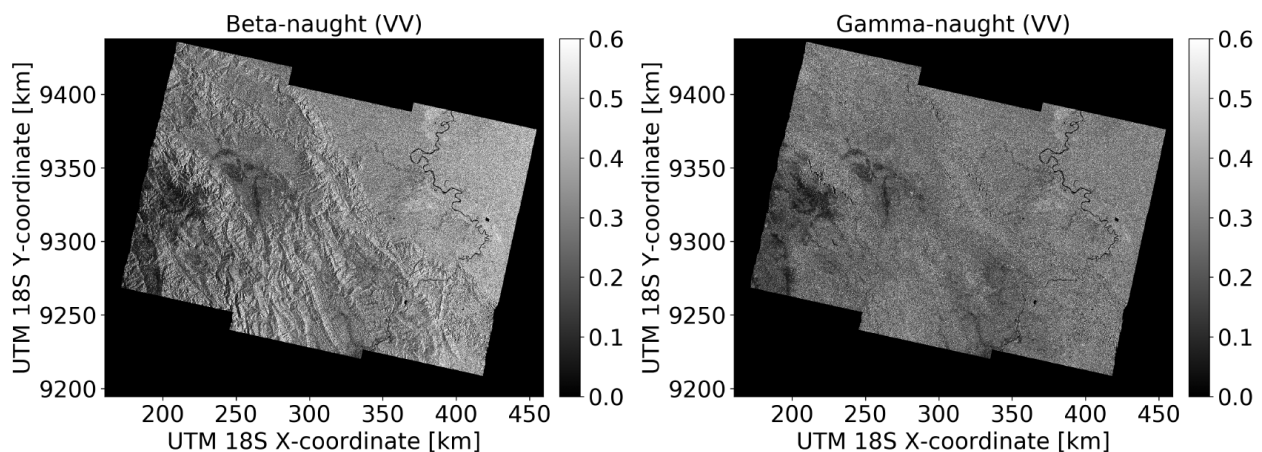


Figure 4.5: Geocoded Sentinel-1B SAR backscatter (VV) without (left) and with (right) radiometric terrain correction (RTC) collected over the Andes in the North of the Department of San Martín (Peru). Both geocoding and RTC steps were performed using the area projection algorithm [Shiroma, 2020]. The RTC reduces the dependency of the SAR backscatter with respect to the topography. Source: [Shiroma, 2020].

Assessment 1 for task 2 - Statistical comparisons: The first approach is a comparison of radiometrically terrain flattened γ^0 values for foreslopes, flat terrain, and back slopes from [Small, 2011]. The relevant validation sites are listed in Table 4.2 as “Flattening” and mapped in Figure 4.6. For each site we create RTC-S1 image stacks in ascending and descending orbit geometry. For each image in these stacks, we are

1. selecting forested areas and masking out water bodies to focus on land surface scattering and avoid biases related to the specular scattering response characteristic for surface water;
2. generating masks demarcating foreslope, flat, and backslope regions within each image;
3. calculating histograms of cross-polarimetric backscatter estimates for each image and separately for foreslopes, flat regions, and backslope areas;
4. Test whether the median radar backscatter of RTC-S1 products over an area of foreslope is within 1dB of the median radar backscatter over an area of backslope for forested land-types, for at least 80% of all validation products considered. A threshold of 1dB was chosen to accommodate for

natural variability of radar backscatter in forested regions and to ensure generation of downstream DSWx products at the required performance.

A figure from [Small, 2011] illustrates the expected empirical distributions of the foreslope (solid), flat (dashed), and backslope (dotted). After RTC, foreslope and backslope histograms are slightly misaligned to flat areas due to the secondary effect of incidence angle, though the misalignment is much less than other methodologies compared in [Small, 2011].

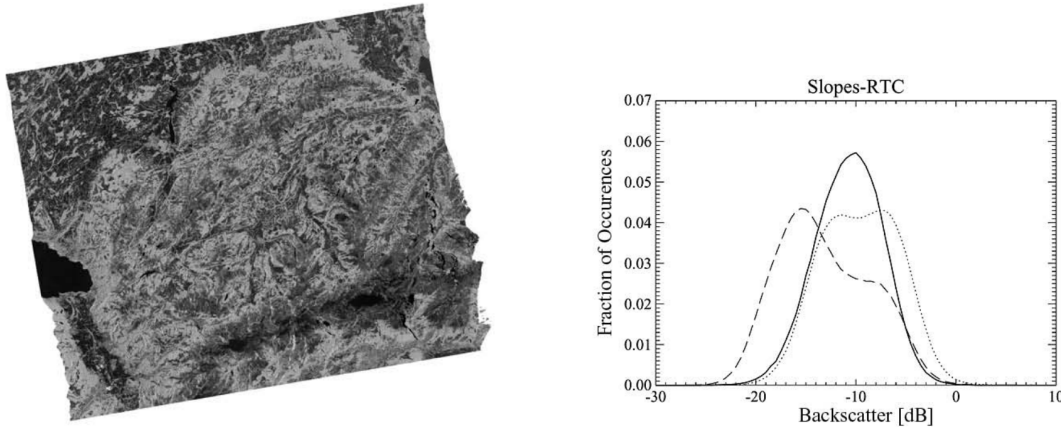


Figure 4.6: From [Small, 2011], left is the RTC image and right is the distribution of backscatter across foreslope (solid), flat (dashed), and backslope (dotted).

Assessment 2 for task 2 - Regression analysis between terrain flattened γ^0 and local incidence angle (θ_l): A regression analysis between terrain flattened RTC-S1 radar backscatter and local incidence angle information can be used to assess terrain flattening performance [Truckenbrodt et al., 2019]. For this, we compare the OPERA RTC-S1 with RTC imagery from ASF processed using GAMMA (see Section 3.4.4). There is no absolute number of quality for terrain flattening other than that GAMMA has shown to achieve the flattest response.

To be able to analyze terrain flattening quality from several images in a time series, it is essential that backscatter properties of corresponding land cover surfaces remain relatively consistent over the period of analysis. Therefore, we are restricting our focus to backscatter values in each image to forested areas, which return more stable backscatter over time and across incidence angles than most other land cover classes [Truckenbrodt et al., 2019]. We utilize a Hansen, et al. forest map from 3.4.6.

As measures of quality of terrain flattening, a linear regression of the form

$$\gamma^{0,OPERA} = a_{\theta,OPERA} + b_{\theta,OPERA} \cdot \theta_{l,OPERA}$$

is derived between terrain flattened $\gamma^{0,OPERA}$ and local incidence angle $\theta_{l,OPERA}$, and the slope $b_{\theta,OPERA}$ of the linear regression and its standard deviation ($\sigma_{b_{\theta,OPERA}}$) are derived. Optimal terrain flattening would yield $b_{\theta,OPERA} = 0$ for areas of uniform scattering for all observed incidence angles, indicating that all slope dependence was removed.

As not all sites will yield $b_{\theta,OPERA} = 0$ (e.g., due to potential variations of DEM quality and surface properties), this analysis is performed as a relative analysis between OPERA RTC-S1 products and RTC data generated with an established RTC workflow at the NASA Alaska Satellite Facility DAAC that is based on the GAMMA software (see Section 3.4.4). Analyzing RTC-S1 imagery relative to ASF RTC products will ensure that potential landscape-related variations of radar brightness will not result in failure

to pass validation as these variations will affect both products equally. It will allow separating algorithm deficiencies from potential residual landscape effects.

For all areas designated as use case “Flattening” in Table 4.2, we will generate RTC-S1 and ASF RTC products from identical input, DEM, and ancillary data. Based on these inputs, we will perform statistical significance tests evaluating if the regression slope $b_{\theta, OPERA}$ is significantly higher than the slope $b_{\theta, ASF}$ derived from the ASF-processed equivalent product ($H_0: b_{\theta, OPERA} \leq b_{\theta, ASF}$) at the 95% confidence level.

The RTC-S1 Validation sites are listed in Table 4.2 and displayed graphically in Figure 4.4. Sites are selected for a variety of tests, detailed in Section 4.1.6, for validating the RTC requirements for geolocation accuracy and radiometric terrain flattening performance.

Site	Reference Granule	Terrain Type	Use Case
Amazon Rainforest, Brazil	S1A_IW_GRDH_1SDV_20150430T092134_20150430T092159_005711_00754A_4706	Agriculture, Hills	Flattening
California, US Ascending	S1A_IW_GRDH_1SDV_20201107T015035_20201107T015100_035136_0419CE_43A4	Urban, Hills, Desert, Corner Reflectors	Geolocation Accuracy
California, US Descending	S1A_IW_GRDH_1SDV_20201107T135216_20201107T135241_035143_041A16_93F1	Urban, Hills, Desert, Corner Reflectors	Geolocation Accuracy
Delta Junction, AK Ascending	S1B_IW_GRDH_1SDV_20200730T032006_20200730T032031_022695_02B12E_BB4A	Hills, Mountains	Flattening, Coregistration
Delta Junction, AK Descending	S1A_IW_GRDH_1SDV_20200716T161220_20200716T161245_033482_03E13C_D6F5	Hills, Mountains	Flattening, Coregistration
Fairbanks, AK Ascending	S1B_IW_GRDH_1SDV_20200730T032031_20200730T032056_022695_02B12E_95A9	Hills, Mountains, Urban	Flattening, Geolocation Accuracy
Fairbanks, AK Descending	S1A_IW_GRDH_1SDV_20200716T161155_20200716T161220_033482_03E13C_7D26	Hills, Mountains, Urban	Flattening, Geolocation Accuracy
Kansas, US	S1A_IW_GRDH_1SDV_20190705T125304_20190705T125333_027982_0328E1_8BA5	Flat, Agriculture	Flattening, Geolocation Accuracy
Mato Grosso, Brazil	S1A_IW_GRDH_1SDV_20150614T090018_20150614T090043_006367_008640_4D51	Rainforest, Hills, Agriculture	Flattening, Coregistration
Vermont / New Hampshire, US Ascending	S1A_IW_SLC_1SDV_20200831T224357_20200831T224430_034157_03F7B6_3237	Forested, Hills	Flattening
Vermont / New Hampshire, US Descending	S1A_IW_SLC_1SDV_20200830T105658_20200830T105726_034135_03F6EB_BB87	Forested, Hills	Flattening
Appalachian Mountains, US	S1A_IW_GRDH_1SDV_20210518T231517_20210518T231542_037949_047A98_00A2	Mountains, Hills	Flattening

Table 4.2: List of 12 RTC S1 Validation Sites; sites selected to represent diverse terrain types.

Proposed Calibration/Validation Reference Scenes

Direction



Esri, FAO, NOAA, USGS, Esri, USGS

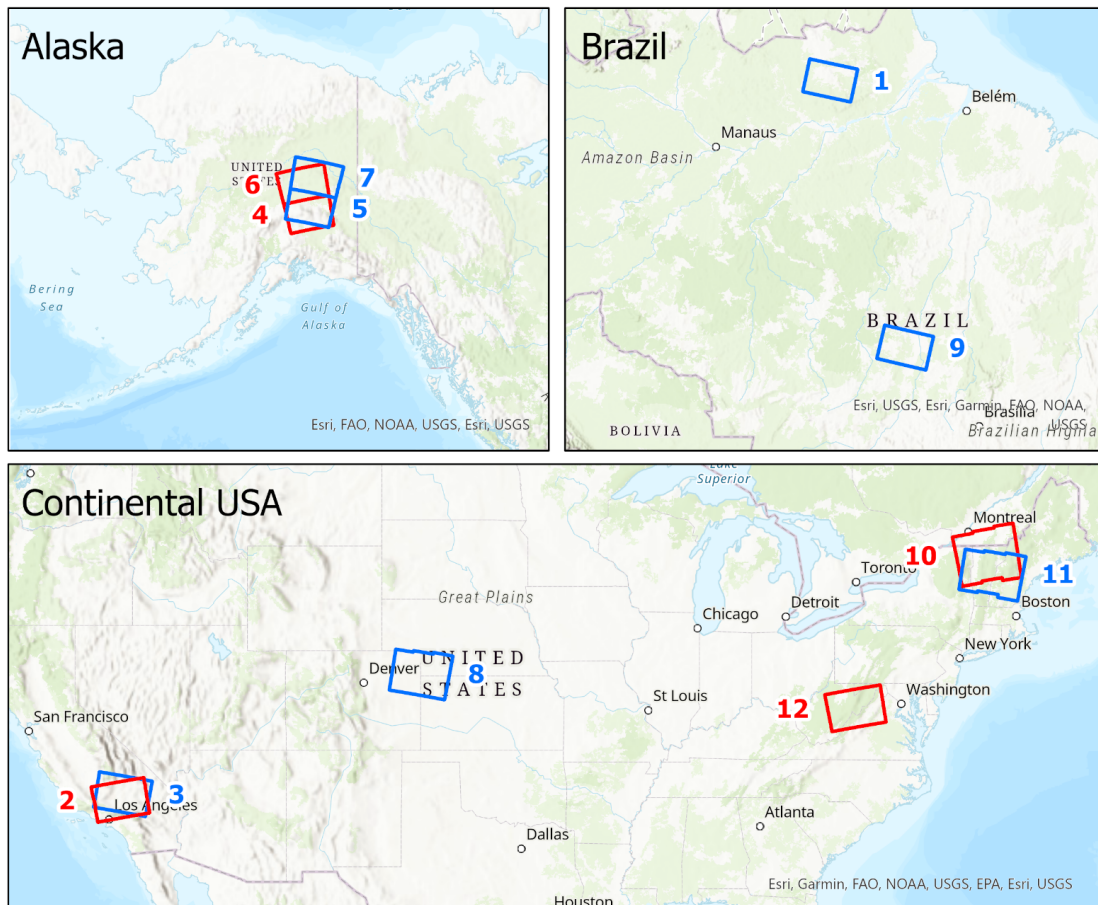


Figure 4.7: Proposed RTC-S1 Validation Sites. Both ascending and descending passes will be utilized where possible. The number next to the frame with the same color indicates the expected number of images that will be available in that frame geometry.

4.4 Displacement Product

The displacement suite of products has two requirements, one for each relevant OPERA product: DISP-S1 and DISP-NI. The two requirements will employ the same validation strategy using the velocity measurements for each of the DISP products. For clarity, the DISP-NI requirement is repeated below:

The NISAR DISP product (DISP-NI) shall measure from a single line-of-sight geometry surface displacement rates with an uncertainty of 5 mm per yr or better over length scales $0.1 \text{ km} < L < 50 \text{ km}$ from 1 year of regularly sampled (i.e. at least 80% of acquisitions from nominal 12 days repeat passes) NISAR data (i.e. from HH polarization imagery from the L-band main band), in regions where the interferometric signal is maintained (i.e., the temporal coherence is above 0.5).

Whereas for DISP-S1:

The Sentinel-1 DISP product (DISP-S1) shall measure from a single line-of-sight geometry surface displacement rates with an uncertainty of 3 mm per yr or better over length scales within $0.1 \text{ km} < L < 50 \text{ km}$ from 4 years of regularly sampled (at least 80% of Sentinel-1 acquisitions at 12 days sampling or better) Sentinel-1 A/B data (i.e. VV polarization imagery in IW-mode), in regions where the interferometric signal is maintained (i.e., the temporal coherence is above 0.5).

We note that the temporal coherence portion of the requirement only *focuses on our validation activities* and will *not* limit the spatial coverage when the OPERA L3 DISP products are generated.

The displacement validation strategies discussed in this document are consistent with those as planned by the NISAR Solid Earth Science team for validating its secular displacement science requirement (see NISAR [Solid Earth ATBD document](#)). More detailed descriptions can be found in [[Lohman and Simmons, 2005](#)], [[Murray, et al. 2019](#)] and [[Weiss et. al., 2020](#)]. Below, we summarize the general procedures in order to implement our validation approaches (i.e. VA1, VA2, VA3) including the two primary statistical analyses that we will perform.

1. *InSAR LOS Velocity estimation:* We will estimate the long term line-of-sight rates and corresponding uncertainties by fitting a time-series function which includes parameters such as constant offset, linear rate, and seasonal terms across the displacement history, derived from the reconstructed displacement time-series using consecutive DISP products in time (*see* [AD3] Figure 3.1C regarding time-series reconstruction). Where needed, we will include additional complexities (e.g. Heaviside step, exponential or logarithmic function for coseismic and postseismic events/processes). Note that the displacement rates are estimated separately for NISAR and Sentinel-1 repeat tracks.
2. *Validation Data LOS Velocity estimation:* We will project the independent Validation displacement data (e.g. GNSS time-series) in the radar look direction and then estimate the velocity over the same period of time and with the same time-series models as from step 1. The GNSS time-series data will have its processing and noise corrections applied (see Table 3.9 for a summary and Section 3.3.4 for a discussion of this data). As an example, the GNSS time-series data from UNR is corrected for instrument changes, ionospheric and tropospheric propagation delays, as well as solid earth tides [[Blewitt et al., 2018](#)].
3. *Statistical analyses:* We will mainly implement similar analyses as exemplified in Figure 4.8 to assess if the OPERA L3 DISP product requirements are achieved.

- Variogram analysis for the line-of-sight velocity residuals will be performed. Using the structure function, values at length scales $0.1 \text{ km} < L < 50 \text{ km}$ will be extracted and then compared to the OPERA L3 DISP requirements for validation.
- We will calculate the mean and the standard deviation of the velocity residuals (e.g. perform scatter plot/RMSE analysis) and then assess if the OPERA L3 DISP requirements are satisfied.

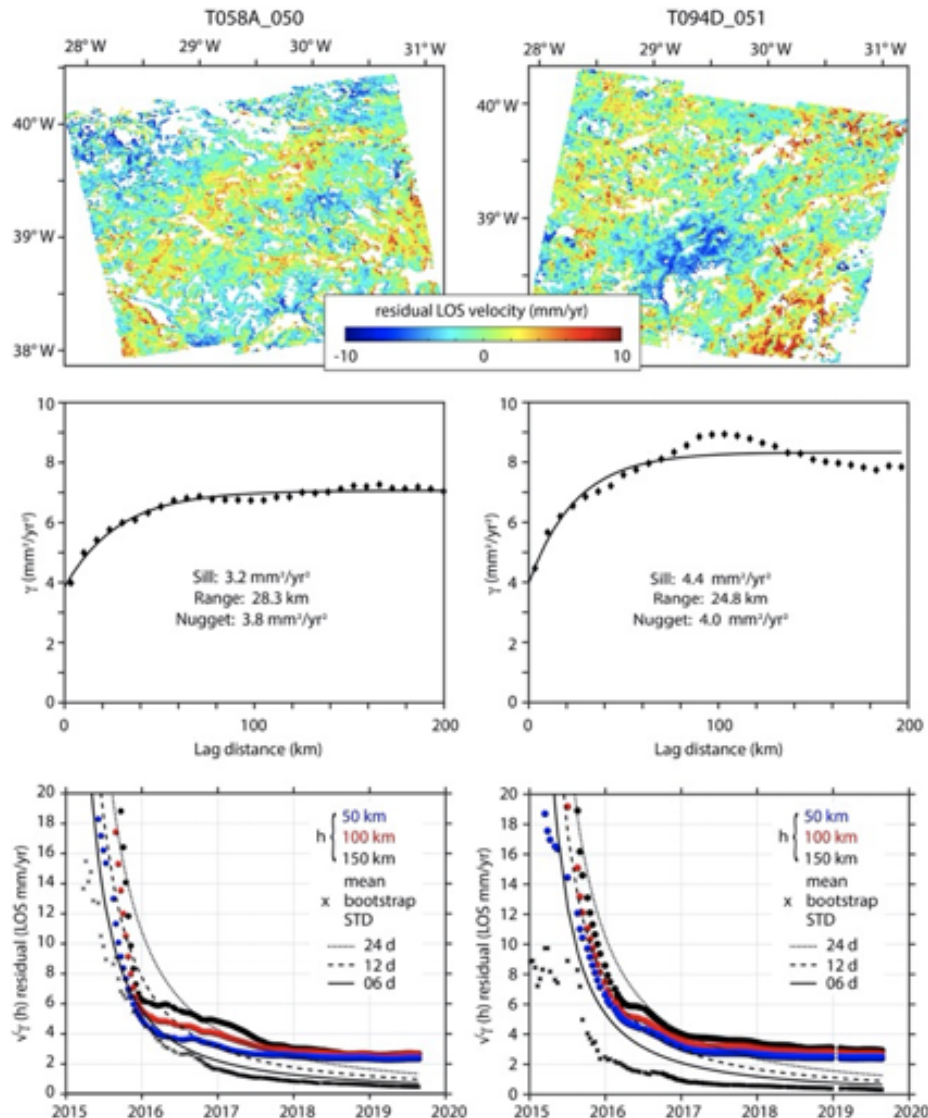


Figure 4.8: From [Weiss et. al., 2020], showing error estimation results for a fault study in the Anatolian region. (Top) Residual LOS velocities for the ascending (left) and descending (right) frames. (Middle) Semi-variograms for the LOS velocities. (Bottom) LOS velocity uncertainty as a function of time for different length scales (e.g. blue: 50 km, red: 100 km, black: 150 km) as well as the theoretical model fits for various acquisition intervals (i.e. 6-, 12-, 24-days).

For all the validation analyses, we will report the accuracy that can be achieved with and without applying the optional correction layers embedded within the L3 DISP products, including the ionospheric propagation delays (e.g., [Fattahi et al., 2017]), tropospheric propagation delays (e.g., [Jolivet et al., 2011] and [Doin et al., 2009]), and solid earth tide effects (e.g. [Agnew 2012]).

Because the OPERA project is generating L3 DISP products that have both Persistent Scatterers (PS) and Distributed Scatterers (DS), we are evaluating the stability of these pixels spatially and temporally during the validation step (e.g. note that each DISP product will have a quality raster layer, *see* [AD3] Figure 3.1). PS are characterized by dominant scatterers within their resolution cells resulting in a very stable phase over time. DS on the other hand are typically natural targets consisting of possibly low coherent reflectivity values but without a single dominant scatterer. DS are the result of the average response of neighboring scatterers within its resolution cell. Recently, studies have shown that DS pixels are subject to phase misclosure issues but does not however apply to PS pixels since phase closure is a consequence of averaging across different resolution cells. Regardless of whether we are correcting for phase misclosure, we will assess the quality of products with respect to the L2 requirements that drive the L3 DISP products. We will calculate the temporal coherence of the pixels for NISAR and Sentinel-1 separately to evaluate the phase stability of DS and PS points over the same area for both the ascending and descending frames. Temporally decorrelated pixels will be masked during validation.

Geographic scope and validation sites: The OPERA DISP product has a North America (NA) geographical scope (Figure 4.9) covering various terrain types and surface scattering. We have selected our displacement validation sites to be representative of the geographical scope for both negligibly deforming and deforming regions in NA (i.e. the latter is due to various natural and anthropogenic deformation processes such as vertical land motion, tectonics, and volcano deformation), with a range of seasonality and surface characteristics including: 1) vegetation cover, 2) rural, urban or mixed environments, and 3) terrain type (i.e. flat or mountainous). Table 4.3 summarizes the proposed validation sites and the target thematic areas and types of deformation that can be captured on these locations. The description of the sites including information on how we are going to validate them are also noted and are further described below (validation approaches). Other areas in North America that are negligibly deforming and have temporally coherent PS and DS pixels will be identified and added by the OPERA's PST. We will utilize the 12-day North America coherence maps that were computed from several years of Sentinel-1 data (see Figure 4.10) as a reference guide to ensure that the selected negligibly deforming validation sites meet the temporal coherence requirement.

Validation approaches: Three validation approaches will be performed separately depending on the Validation dataset that will be used to assess the DISP product requirements. Each validation approach is dubbed in this document as Validation Approach n (VAn) for $n = 1, 2, 3$, i.e. VA1, VA2, and VA3. The validation approaches are primarily differentiated based on the surface displacement occurring at the validation site (i.e. deforming or negligibly deforming) and/or the instrument used to validate the InSAR measurement (i.e. GNSS, corner reflectors, InSAR).

Validation approach 1 (VA1): Direct comparison of GNSS and InSAR is performed over deforming regions with dense GNSS networks including California, Hawaii, Houston/Galveston, and Unimak Island. More validation sites including over Mexico may be added in the next version of this document based on station density analysis necessary to perform the validation of the DISP requirement. These validation sites are shown in Figure 4.9 and detailed in Table 4.3. Temporally decorrelated pixels on the InSAR data (i.e. temporal coherence ≤ 0.5) are removed and a water mask is applied during validation. An initial inspection of the GNSS time-series and the study area is done to select a stable reference area for all the InSAR data (i.e. reference point noise analysis). We will perform root-mean-square error (RMSE) analysis (i.e. statistical analysis 3b) to assess if we meet our DISP product requirements.

Validation approach 2 (VA2): InSAR residual analysis is conducted over a variety of negligibly deforming regions or in areas that are located outside the deformation field within the SAR footprint. For negligibly deforming regions, we target those that have maintained coherence (i.e. temporal coherence > 0.5) PS and DS pixels within a frame. For a scene that includes a deformation field, we will mask first the deformation area as well as those that have temporal coherence ≤ 0.5 prior to residual analysis.

The InSAR displacement rate residuals will be computed from a dense grid of randomly selected pairs of pixels to evaluate the accuracy as a function of distance. Since the selection of pixels is unbiased (or random), the inherent accuracy is evaluated not only on PS pixels but on a combined performance of both PS and DS within the SAR footprint. Note that this VA2 strategy supports and complements VA1 (e.g. length scales, $L < 0.1$ km) and also removes the limitation in terms of the number of GNSS stations that can be used to validate the product requirements. We will perform a variogram analysis as described above (i.e. statistical analysis 3a) to assess if we meet the DISP requirements (i.e. 3 mm per yr or better for Sentinel-1 and 5 mm per yr or better for NISAR, over length scales of $0.1 \text{ km} < L < 50 \text{ km}$).

Validation approach 3 (VA3): In addition to the two validation approaches described above, we are also performing a validation experiment using the corner reflectors that we will deploy across the creeping segment of the San Andreas Fault. The precise height and locations of the CRs are obtained from their co-located GNSS stations or by conducting geodetic surveys annually or semi-annually. We will take the 3-component CR position values and project them to the satellite's line-of-sight direction to estimate the CR LOS velocities. We will then calculate the residuals by directly comparing the CR LOS velocities and InSAR LOS displacement rates. To test if the requirements are achieved for NISAR and Sentinel-1, we will implement statistical analysis 3b as described above. Additional rigorous statistical modeling and tests will be considered for additional verification of the requirements.

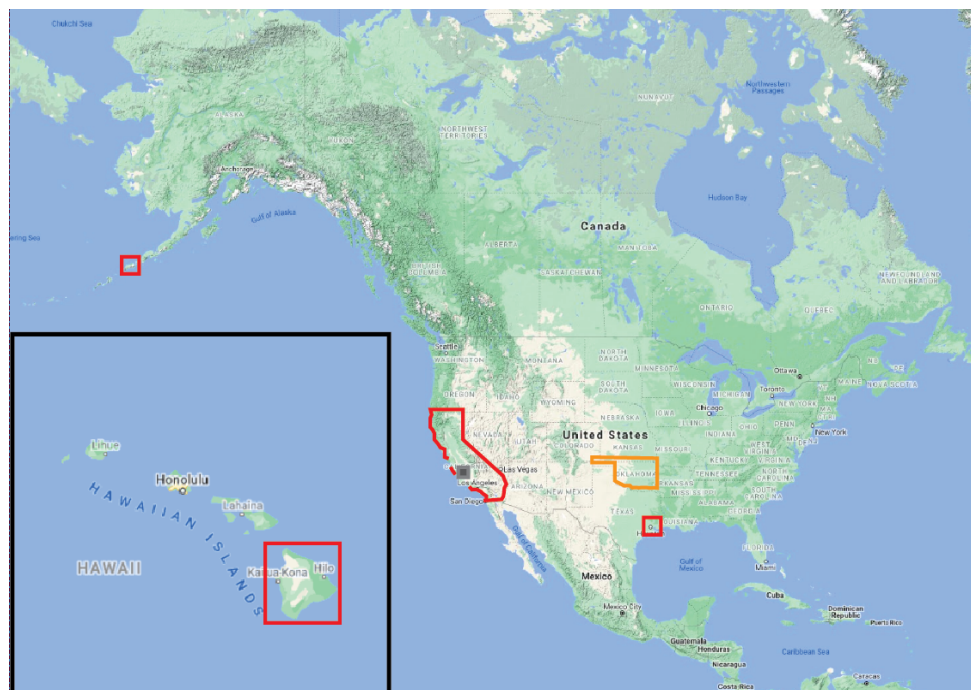


Figure 4.9: Proposed Validation displacement sites from Table 4.3 displayed geographically. Red outlines are regions which have dense GNSS networks and are considered deforming. Orange outline (Oklahoma) is considered a negligibly deforming region with dense GNSS stations. Gray square in California is the target site of the CRs to be deployed by OPERA.

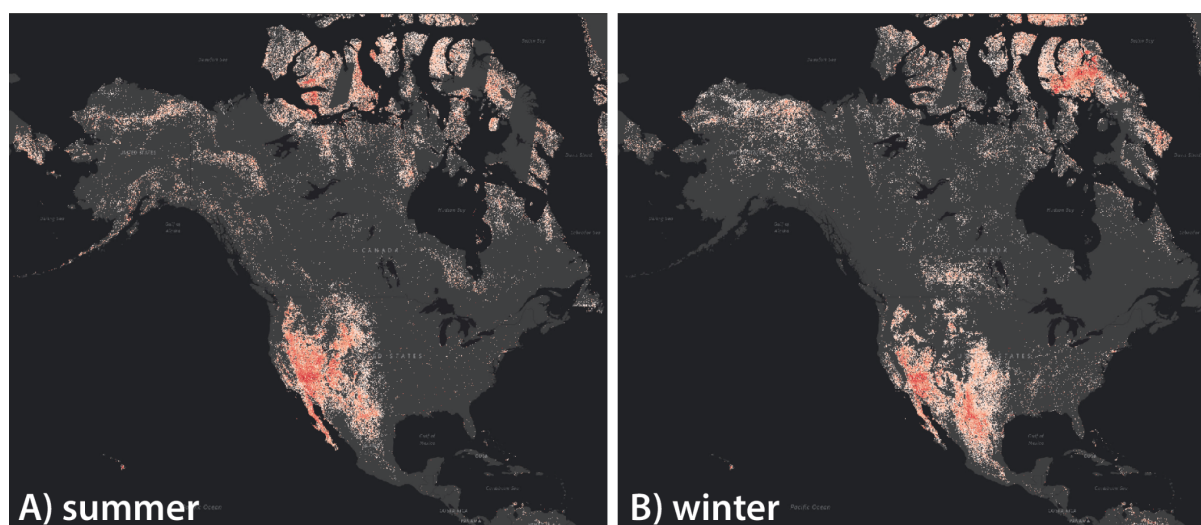


Figure 4.10: Temporal coherence map (i.e. only temporal coherence > 0.5 is shown, towards red means higher coherence) calculated from 12-day Sentinel-1A/B data since 2014 over North America for the **A)** summer and **B)** winter seasons. Notice the loss of coherence in some areas mainly due to vegetation growth and snow cover during the summer and winter periods, respectively. These maps including the fall and spring coherence data will be used as reference guides to ensure that the selected OPERA CSLC and DISP CalVal sites meet the coherence requirements that are only specified for the validation activities.

Target Thematic Area	Type of Deformation	Validation Sites	Validation Approach	Climate	Characteristics	PS/DS
Tectonic deformation, Volcano deformation, Local Land Subsidence (natural and anthropogenic)	secular, co-seismic, transient	California	VA1, VA2, VA3	Temperate/Dry/ Hot	Mix of rural and urban areas, range of relief and decorrelation sources, change in base elevation *100+ cGNSS available **CRs to be deployed by OPERA Validation thru JPL	e.g. 1 CR is planned to be deployed on an agricultural field in Central Valley
Volcano deformation	transient	Big Island, Hawaii	VA1, VA2	Tropical/Rainforest, Tropical/Savanna	Mostly urban, rain forest, relief, tropical climate, island, lava flows, decorrelation source, volcano inflation/deflation *58 cGNSS available	PS+DS in topo-correlated area
Stable deformation	stable	Oklahoma	VA1, VA2, VA3	Temperate/Dry/ Hot	Agriculture, strong atmosphere, no relief *18 cGNSS	To leverage NISAR CRs

					10 CRs deployed by NISAR thru NASA/Univ Oklahoma	
Local land subsidence (anthropogenic and natural sources)	secular, transient	Houston-Galveston, Texas	VA1, VA2	Subtropical/ Wet/ Hot	Mostly urban, swamp, hydrocarbon fields, sinkhole formations, salt dome, no relief, Southern US climate *NA	PS+DS in a mix of urban and swap environment
Volcano deformation, Tectonic deformation	secular,transient	Unimak Island, Alaska	VA1, VA2	Cold/Wet	Subarctic, ocean island, unstable atmosphere, range of relief, snow *14 cGNSS available	TBD

Table 4.3: Proposed Validation displacement sites chosen to represent a diversity of deformation processes in Northern America over various climate and terrain locations.

4.5 Dynamic Surface Water Extent Product

There are currently three categories of Validation that will support the Dynamic Surface Water Extent (DSWx) product suite and validation of its requirements introduced in Section 2.1:

- Validation for open water detection across the DSWx suite i.e. requirements 2 and 4
- Validation for partial surface water detection for DSWx-HLS, i.e. requirement 3
- Validation for inundation under vegetation detection for the DSWx-SAR suite, i.e. requirement 5

The DSWx products shall achieve a pixel classification accuracy of at least 80% for “open water”; the DSWx-HLS product shall receive 70% for “partial surface water”; this accuracy is TBD for “inundated vegetation” with respect to the DSWx products derived from SAR. More specifically, the ADT will perform a trade study to determine the feasibility and performance baseline of the DSWx-S1, DSWx-N, and DSWx-SWOT for inundation vegetation detection. This will be clarified in future versions of this document.

For each category of validation, there are three prongs that all will share: 1) selection of the validation sites, 2) generation of the validation dataset to compare against the provisional DSWx product, and 3) the final comparison between the Validation data and the OPERA DSWx product. We will carefully detail the evaluation of open water detection and then be able to adapt this description for the other two categories of validation activities for the DSWx suite.

Validation of Open Water: We first outline validation of the open water accuracy requirement that will be shared across all DSWx products. This approach is derived from [Jones 2019] and [Pickens, et al. 2020]. Although we follow the general approach, particularly of how we determine validation sites and the use of high resolution commercial satellite data (e.g. PlanetScope and Worldview from Section 3.4.1), we

utilize additional data sources derived from the field work (Section 3.2.2) to generate more expansive validation data.

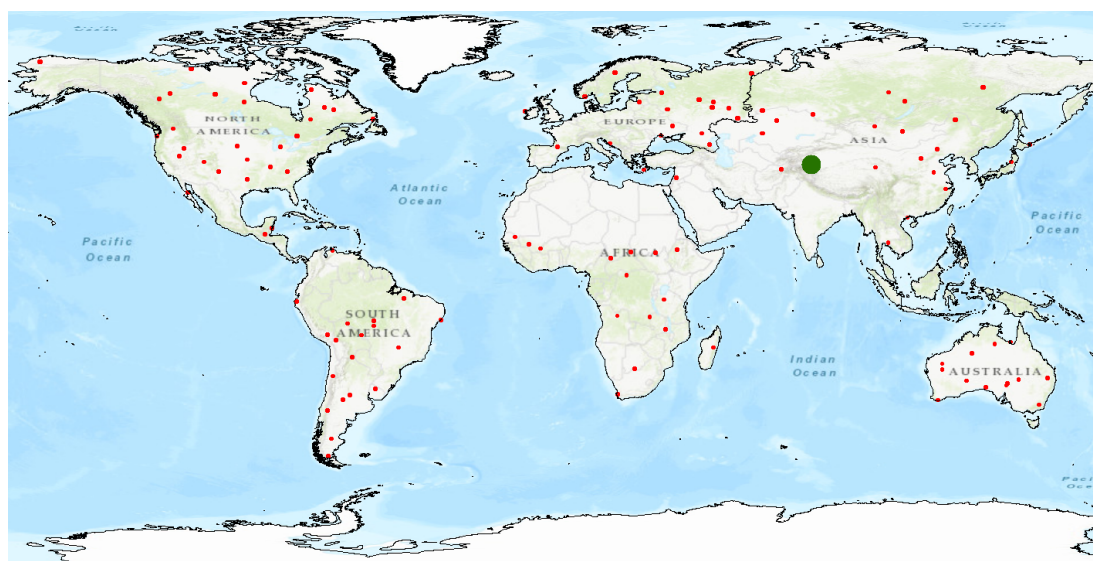
First, we discuss how validation sites are selected following [Pickens et al., 2020]. Pickens et. al. used preliminary estimates of water extent estimated from Landsat series to partition 20 km square grid cells covering the global land surface into water percentage strata, i.e. populations of low/medium/high water presence (specifically, 0–0.08%, 0.08%–2%, and >2%, respectively). Such stratified sampling improves efficiency and accuracy of DSWx Validation over Validation based on random or systematic sampling because inland water accounts for only 3–4% of the total inland area [Pickens et al., 2020]. With these grid cells, Pickens et al. select high resolution satellite imagery chips falling within each of these different strata for their Validation activities. The high-resolution imagery does not cover an entire 20 km grid cell, but these strata provide efficient ways for sampling the corpus of high resolution satellite imagery.

Now, the aforementioned stratification and subsequent validation was based on data that predates DSWx image inputs by 5 years or more. Therefore, we slightly modify their approach to select imagery relevant for DSWx. We consider the random strata from [Pickens et al., 2020] (see Figure 4.11) and derive 10 high resolution optical chips from each strata. First, we select 10 high-resolution imagery from the Pickens et al. strata at sites where these high-resolution commercial satellites and DSWx images are coincident in both space and time (relative to expected dynamics of surface water extent). Then, we classify each high-resolution image by a combination of expert analysis and automated classification and use this classification to determine the updated strata relevant for our validation activities. From these new strata, we will identify a minimum of 30 high-resolution optical chips, that is 10 from each new re-classified strata using the methodology above.

The commercial data will be processed using combined automated classification and visual inspection to generate high-resolution open water masks. As an example, similar Validation activities were used for the USGS DSWE product [USGS link]. In Figure 4.12, we illustrate how this comparison is made from previous Validation activities related to the USGS DSWE product. Here, the DSWE product is compared with a RapidEye water mask derived by Pickens et. al. (2020). For this Validation portion of activities, only binary masks will be compared. Note that Figure 4.12 is meant only as a guiding example as we will only be validating “open water” against “not open water” for this validation activity.

From these high-resolution open water maps, we will validate the DSWx open water classes. Specifically, we will aggregate (i.e. reproject via interpolation) the high resolution data into the DSWx reference frame and consider the pixels from the high resolution mask that are entirely covered by water. From this, we will evaluate the classification requirement for the DSWx-suite.

Now, above is only a slight modification from the validation approaches detailed in [Jones, 2019] and [Pickens et al., 2020]. However, in addition to high resolution satellite imagery, we will utilize numerous data sources described in Section 3 to both improve the high-resolution open water maps and generate auxiliary maps. In addition to Planetscope, we will use MAXAR (Worldview) and UAS imagery (if available over the calval sites, though not required), previously published USGS data, and possibly previously published NASA Validation data. Furthermore, we will select high resolution chips over the field work sites and use the field work to augment the open water map derived from the high-resolution optical imagery (see Section 3.2).



Sources: Esri, HERE, Garmin, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, Geobase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), Swisstopo, © OpenStreetMap contributors, and the GIS User Community

Figure 4.11 : Locations of Validation sample sites used by Pickens et. al. (2020). The site highlighted with a larger green symbol is that portrayed in Figure 4.12 (next figure). Water masks for a subset of these sites will be updated at locations where coincident high-resolution commercial imagery and DSWx image inputs become available.

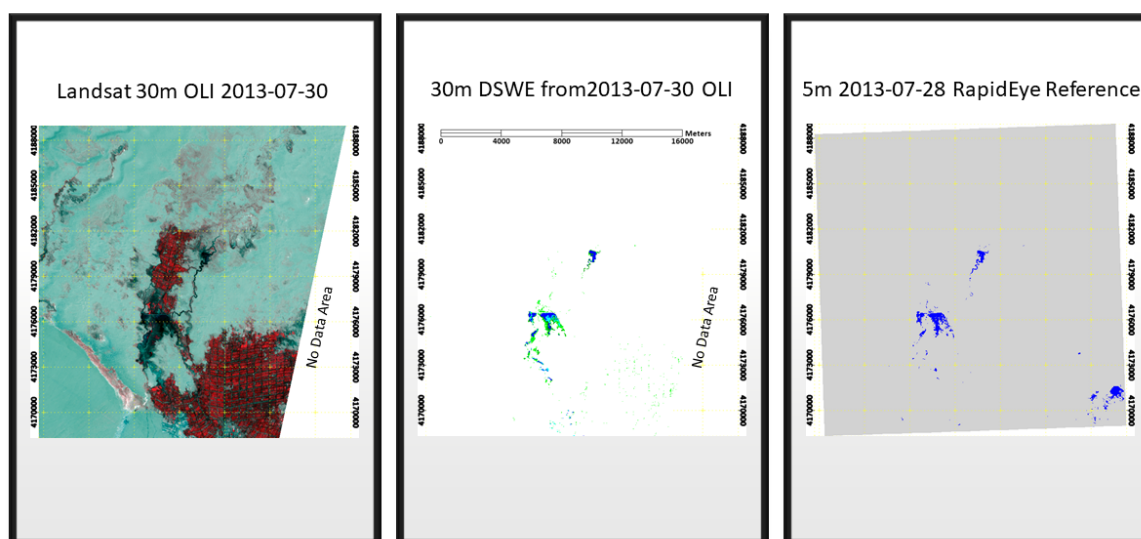


Figure 4.12: Example of Landsat OLI data (Left), USGS DSWE output (Center) and RapidEye-based water mask (Right) for 1 published 5 km by 5 km sample site (Pickens et. al 2020) located in western China near the border with Tajikistan (large dot in Figure 4.11). Note that the complete USGS DSWE product classification output is shown above.

Validation of Partial Surface Water: the validation of partial surface water for DSWx-HLS will follow the same approach as for “open water”, after obtaining high-resolution optical chips from 3 different re-classified strata and generating water extent maps. We use the high-resolution open water maps from the previous activity because reprojection of the high-resolution maps into the lower resolution DSWx-HLS

frame yields pixels that are partially inundated, i.e. having an inundation percent between 0.5 and 1. This will provide one assessment of the partial surface water class seen in DSWx-HLS. The lower left image in Figure 4.13 (taken from [Pickens, et al. 2020]) portrays as dark tones the areas that may be identified as ‘open water’ in high-resolution but may constitute mixed or partial surface water pixels at Landsat resolution shown in the upper right image.

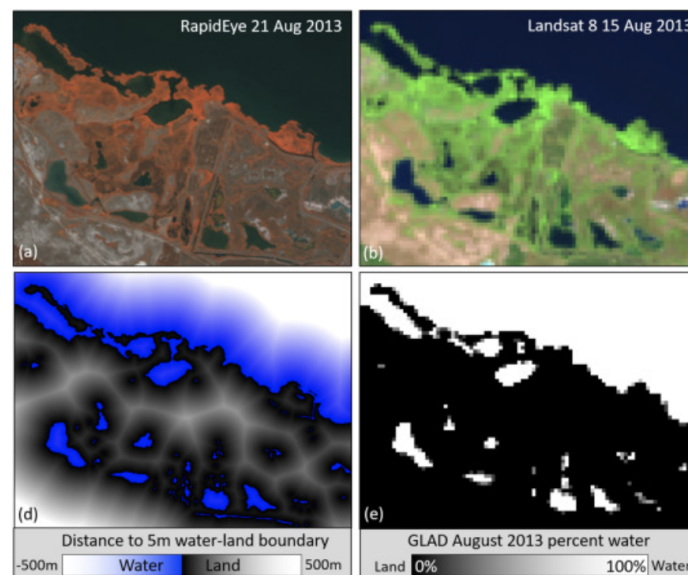


Figure 4.13 : From [Pickens, et al. 2020]. Left column: high resolution optical imagery (top) and open water mask (bottom). Right column: Landsat imagery coincident with high-resolution image (top) and processed to surface water (bottom).

However, there will be additional modes of assessment for this more challenging detection. Where vegetation cover is particularly prevalent (e.g., marsh, wet prairie, and swamp environments), it is difficult to rely on classification maps generated from high-resolution optical data alone (e.g. PlanetScope and MAXAR) for validation purposes. For this reason, resource network data (i.e. EDEN and USGS gauge data), high-resolution image classifications *trained and validated using OPERA-collected field data*, and the field data themselves, such as wetted perimeters around partial surface water areas at Landsat scale, will be used to evaluate DSWx partial surface water and water under vegetation classes. OPERA field data will be invaluable for this Validation activity because accurate inundation vegetation mapping at the resolution and at the scale required is rare. In this case, we will have high resolution maps and transects that will both provide validation sources to compare the DSWx-HLS products (see 3.2.2 for a description of the field work sites). The former will be processed for validation in precisely the same way as the globally dispersed, high-resolution commercial image chips (reprojecting the data in the DSWx-HLS frame and comparing relevant pixels). Finally, transects at DSWx field sites will provide higher temporal sampling across numerous DSWx-HLS products and whether or not suitable high-resolution data can be leveraged via the transects, transect point data can be used to determine partial surface water presence within associated pixels. An example of this analysis is shown in Figure 4.14 taken from [Jones 2015, Jones 2019], where the “agreement rate” is the accuracy of the class compared against in-situ measured stage and elevation-derived declarations of ‘water’ or ‘not water’ through time for water extent product pixels that are coincident with gauged-based measurement.

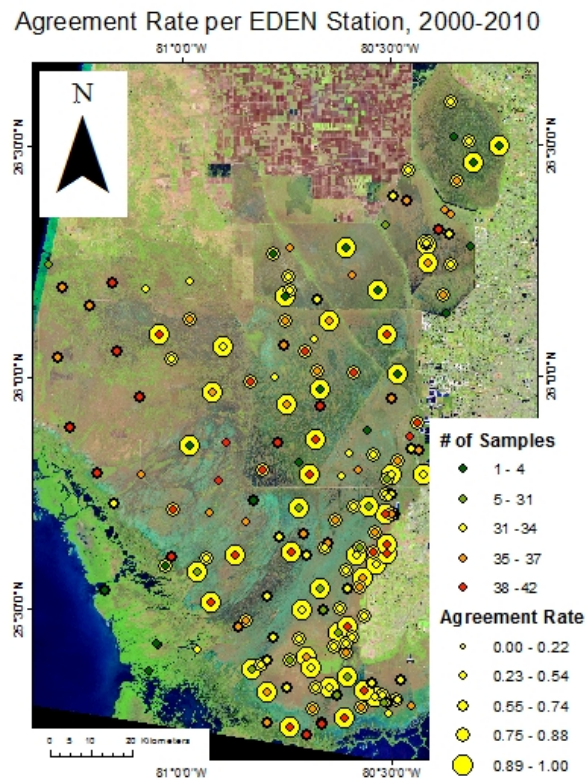


Figure 4.14 : From [Jones 2015]. Bivariate map indicating sample size and agreement rate with water level as measured at EDEN gauge sites and ‘inundation’ as detected by a prototype DSWE subpixel inundation detection algorithm. Comparisons were made between gauges and “clear” pixels over a 10-year period.

The partial surface water class accuracy assessments described above will provide a range of biomes in which to assess performance. However, the analysis conducted in Florida (e.g., Figure 4.14) will be repeated to assess uncertainty associated with the specific implementation of revised DSWE to HLS and can be repeated over time for other DSWx products as they enter validation phases. In the case of the Everglades, vegetation community and local topography providing the context for each gauge were characterized through field survey for 175 of the nearly 300 EDEN gauge locations [EDEN Gauge Metadata]. Locations will be selected across DSWx field sites that provide a range of size and wetland classes where vegetation composition and structure information (percent cover and Leaf Area Index or LAI) will be compiled using LiDAR datasets (where available) in combination with transect and area based sampling of LAI and vegetation cover estimated using AccuPAR LP-80 (Figure 4.15), moose horn densitometers, and hemispherical photography. The field regions are listed in Section 3.2.2. These data will clarify the impacts of vegetation cover on partial surface water and inundated vegetation detection.

Validation of Inundated Vegetation: The approach for measuring inundated vegetation for the SAR derived DSWx product will mirror the approach for the validation of the partial surface water of DSWx-HLS products. The primary difference is that field data will be employed that is specifically selected for determining the sensitivity of the SAR products and is not expected to be detectable from the optical inputs. For example, inundated areas in forests and dense floating vegetation during “leaf-on” seasons (see section 3.2.2). Again, additional analysis relating the vegetation density will be done to determine the impact of the detection as it relates to vegetation. Furthermore, this analysis will specify a performance baseline for inundated vegetation detection rates from SAR sensors.



Figure 4.15: Transect-based data collection using an Accupar LP-80 to generate satellite-product relevant [Jones 2011] vegetation cover information for inundated Everglades wet prairie.

For all the assessments, additional statistical techniques (e.g cross-validation, confusion matrices, precision-recall curves, and error analysis as in [Jones 2015]) will be performed and published.

4.6 Disturbance Product

Validation of DIST product will have two important categories of activities:

1. Validation of the disturbance detection layers in the DIST_ALERT and DIST_ANN products using high-resolution derived disturbance data.
2. Assessment activities related to the current vegetation cover indicator within the DIST_ALERT product.

The first category (described in 4.6.1) is within the scope of the requirements and will be used to determine the accuracy of the vegetation disturbance (see [AD3]). The second category (described in 4.6.2) is related to reporting the statistical relationship of the intermediate vegetation cover indicator layer used by the disturbance algorithm and in-situ vegetation cover as determined by field work.

4.6.1 Validation Strategy for Disturbance Detection in DIST suite

To validate both the DIST_ALERT and the DIST_ANN we employ a stratified random sample of reference data. The global population of all 30m pixels aligned with the HLS pixel grid will be stratified based on disturbance presence and an equal area sample of 30m pixels will be selected. For each pixel randomly selected as a validation site, 3m PlanetScope data will be employed to create reference data of disturbance status and the initial date of disturbance events, where applicable. This reference data will then be compared to the selected pixel in the DIST_ALERT time-series and the DIST_ANN product to ensure the requirement is satisfied.

The reference data for each validation site is created by manually labeling disturbance if necessary for any selected 30m HLS pixel footprint from a time-series of high-resolution Planetscope data. An analyst marks all time steps of the high-resolution data that have a $\geq 50\%$ loss of vegetation cover, resulting in a yes/no reference time-series of $\geq 50\%$ vegetation cover loss. The analyst will also mark the initial date of any discernible loss events $< 50\%$ or whether there was no-disturbance.

For vegetation disturbance validation, per the requirements, the vegetation disturbances with anomalies $\geq 50\%$ mapped in the DIST_ALERT will be considered disturbed and all other land observations will be considered no-disturbance. As the DIST_ALERT product is released at the cadence of the HLS input dataset, the accuracy of the DIST_ALERT product is calculated from all HLS time-steps with respect to the reference disturbance time-series so that the reported accuracy will apply to all observations and thus all phenological stages. The DIST_ALERT product will be compared to the reference time-series by matching the respective dates. There is an associated date for each disturbance/no-disturbance label in the reference time-series which corresponds to the time of the high-resolution Planet acquisitions. For each HLS acquisition date, DIST_ALERT will be compared with the reference data label from the same date when possible. For HLS dates without a coincident reference label, if the closest preceding and following reference date have the same label, this is compared with the DIST_ALERT label. If the two reference dates have different labels this time step is excluded as it is unknown when the disturbance occurred between these two dates. Given that the validation assessment is through an entire one-year period with all HLS acquisition dates evaluated, all seasons are assessed and any errors due to intra-annual variation will be quantified.

We also note that the DIST_ALERT has two types of disturbances: provisional and confirmed. Both types of disturbances in DIST_ALERT will be evaluated against the high-resolution derived time-series. Typically, the confirmed disturbances in the DIST_ALERT product are expected to have a higher accuracy as they have been consistently observed but given that the DIST_ALERT products may be used immediately as the HLS scene is released and *before* a disturbance can be confirmed from repeat observations. Therefore, this broader comparison of both disturbance types in the DIST_ALERT is essential for users to understand. The overall user's and producer's accuracies of both types of disturbances will be reported.

The DIST_ANN product will only be validated against confirmed changes from the previous calendar years⁵. Each date will not be compared for DIST_ANN but rather only the labels of disturbance $\geq 50\%$. The overall, user's and producer's accuracies will be reported.

Figure 4.16 provides a simple example of the comparison through two images of a high-resolution time-series and a Landsat sequence. The first two panels represent two images from the high-resolution time-series. The red box is the single pixel that will be hand labeled. The next two panels illustrate HLS scenes from coincident dates. The rightmost panel shows the disturbance detected by the DIST product and the hand-labeled reference data will be compared at the relevant HLS pixel.



Figure 4.16: Above an example of DIST detection (from [Ying et al. 2017]). The left two panels are two high resolution Google Earth images over the same area in Antalya, Turkey. The next two panels are Landsat acquisitions of the same area. The rightmost panel shows in yellow the vegetation cover loss

⁵ The number of years of previous years compared against will depend on the availability of HLS data.

mapped with Landsat in [Ying et al. 2017]. Note the red box collocated in the center of all images shows that a previously vegetated area was covered by greenhouses and is detected as disturbance.

To ensure the validation meets the requirement criteria through a broad range of sites, we employ a rigorous stratified sampling scheme targeting locations of disturbance, similar to [Ying et al. 2017] and [Pickens, et al. 2020]. Using one year of DIST_ALERT data, we can stratify the global population of 30m HLS pixel footprints into five strata according to different types of observed disturbance categories:

1. Confirmed disturbances (in both DIST_ALERT and DIST_ANN, by definition) with maximum anomaly $\geq 50\%$,
2. Provisional disturbances that are not confirmed (at any point within the year as detected by DIST_ALERT) with maximum anomaly $\geq 50\%$,
3. Confirmed and provisional disturbances with maximum anomaly $< 50\%$ (at any point within the year as detected by DIST_ALERT),
4. Spectral anomalies recorded in the *generic disturbance maximum anomaly* layer of DIST_ANN above a certain threshold⁶ excluding the above strata,
5. No detected disturbance.

From each strata of pixels, we will randomly select 50 HLS pixel footprints to evaluate an annual time-series as described above. Given these strata are global in scope, the sample will include pixels from both low and high latitudes in proportion to their area and the magnitude of omission and commission is statistically quantified within the validation. The entire year will be assessed with high-resolution imagery, evaluating the DIST products for both low and high sun angles.

It has been shown that such stratification is most efficient and most accurate for Validation activities compared to systematic sampling and random sampling of sites [Broich et al., 2009]. Furthermore, without a stratification, the likelihood of evaluating disturbance in a statistically robust manner is not possible because areas of disturbance on annual timescales represent a negligible fraction of the global land surface. In [Ying et al. 2017], it was found that less than 8,000 km² of land experienced semi-permanent loss of vegetation on an annual basis. This area represents less than 0.002% of the global land surface. Thus, to robustly validate the disturbance map requires the stratification derived from a prototype disturbance map.

Map accuracy and unbiased estimates of land cover change will follow [Olofsson et al., 2014] and include a confusion matrix for our nominal definition of $\geq 50\%$ vegetation loss per pixel, as well as standard metrics for different thresholds of change (over, user's, and producer's accuracies, etc.).

4.6.2 Assessment Strategy of the Vegetation Cover Indicator Layer

The assessment of the vegetation cover indicator layer pertains to the vegetation layers of the DIST suite used as inputs to the disturbance detection algorithms and distributed with the DIST product. Although this layer is without a formal requirement, providing a general assessment of this intermediate layer's utility is valuable for users and for increasing the transparency of the disturbance detection algorithm. The assessment uses the 3 by 3 grids of vegetation cover derived from field work described in Section 3.2.3. Each grid is aligned with the DIST_ALERT product pixels and is produced using sub-meter maps from field activities. The field data is associated with a given date and is compared to the DIST_ALERT from the same day as the field work. Specifically, the vegetation indicator layer [AD3] is compared to the co-located 3 by 3 grid of vegetation cover derived from the field work collected during different seasons. A

⁶ Precise threshold to be determined during calibration and detailed in the future ATBD.

comparison of the vegetation indicator layer with respect to field reference maps will be reported, including RMSE and R^2 .

4.7 Ongoing validation of OPERA products

We note that there is a project requirement for ongoing validation activities during the 3-year production phase of this project to ensure that the generated products satisfy the L2 requirements. The frequency of this validation will be clarified in future documents and will depend on the specific OPERA product being validated.

For the DIST-HLS products, the validation strategy robustly determines the validation accuracy based on the spectral characteristics of the HLS imagery. For a given set of spectral signatures over time, the product will be assigned a disturbance class (e.g. provisional or confirmed change). Initial validation of these products will be done using a full year of HLS data and the input HLS reflectance data will be consistently processed throughout the duration of the OPERA project, reducing the need for ongoing validation. However, post-operational validation of DIST_HLS products may be necessary if there is ever any degradation or change in current HLS inputs. Additionally, a robust program of automated verification of the DIST-HLS products will be implemented by the algorithm development team and designed to capture anomalies, systematic errors, and artifacts that could negatively impact product accuracy.

Since the S1 RTC product is a correction based on the geometry of the observation and the digital elevation model of the Earth, after initial validation has shown that the product has the expected characteristics, ongoing validation will not be necessary unless input properties have changed, such as incorporating a new DEM.

DSWx will undergo 4 separate releases as inputs from different sensing systems are accommodated. The staggered launch and operationalization of these systems makes separate DSWx releases necessary. The dynamic nature of surface water extent dictates that new validation products be created to match the collection timeframe of the input source being evaluated. Therefore, for each DSWx product release, a validation report will be produced before the end of each year of operational product generation. Each report will document product accuracy using updated data for the set of validation sites used for initial release 1 product validation. In addition, a small number of sites will be added to the validation dataset with each release. These additional sites may be chosen to (1) re-examine classification accuracy of the previous product deliveries and confirm continued compliance with requirements, (2) examine DSWx product performance given particular challenges, (3) examine DSWx performance given particular user application interests, and/or (4) for sites not yet processed from the original sampling frame to increase the overall statistical power of the uncertainty assessment as resources allow. For example, these additional sites may include locations and time periods surrounding the ice breakup of large water bodies, where the expected accuracy of the DSWx-HLS products may have larger uncertainties. As another example, a site may be chosen given its importance to an issue of resource management, such as stream water allocation to support multiple uses. Thus, ongoing validation activities will become part of the development and application activities for new DWSx products from additional sensors (such as Sentinel-1, NISAR, SWOT) and any created harmonized products.

After primary validation is completed for S1 Coreg SLC and displacement products, during the next product validation period for NISAR Coreg SLC and displacement, a small subset of validation sites will also be used to re-examine accuracy of the previous delivery and confirm continued compliance with requirements.

4.8 Validation deliverables

4.8.1 Validation products

Above, the validation strategies detail how Validation data and OPERA products are compared to ensure satisfaction of the L2 requirements [AD2]. Below, in Table 4.4, all the relevant data and subsequent analysis products are summarized. These input data products and analysis products will be made freely accessible via NASA DAACs (however, there are distribution restrictions on proprietary Validation input data such as from Planet). In some cases, additional avenues of distribution will be included due to inter-agency partnerships.

OPERA Product	Validation Products	Description/details	Availability
L2 CSLC	- Corner reflectors positions and orientation - CSLC offset map (ascending and descending) with radar and geo coordinate values	- Absolute geolocation: CR ID, latitude, longitude, height, orientation - Relative geolocation: CLSC offset map with radar and geo coordinate values	NASA DAACs
L2 RTC	- OPERA RTC products over validation sites - ASF's RTC products (produced with GAMMA) over validation sites - Standard error statistics dictated by comparison	- The RTC products will include basic SAR geometric information for additional comparisons	NASA DAACs
L3 DISP	- Deformation rate time series - OPERA DISP stack over sites	-InSAR: track, frame, asc/des, latitude, longitude, displacement rates, duration/date for rate estimation, corrections applied: atmosphere, ionosphere, and solid earth tides, mask areas -GNSS: GNSS station IDs, latitude, longitude, displacement rates, duration/date for rate estimation, repository (i.e. UNR) -GNSS station pairs: GNSS station1 ID, GNSS station2 ID, length scale/distance between station1 and station2, displacement rate (with respect to station1), duration/date for rate estimation -CR: CR ID, latitude, longitude, height, orientation	NASA DAACs
L3 DSW _x	- Open water extent validation products including a) a binary water mask from sample data and b) a full classification denoting partially inundated pixels - EDEN Water depth maps covering validation datasets - Wetland perimeters in a compatible geospatial format - Presence/absence of surface water at points in compatible geospatial format derived from field work - Standard classification error analysis (precision-recall, confusion matrix)	- Utilize previously published datasets from Sleeter et. al 2021, Pickens et. al. 2019, Jones 2015, Pengra et. al. 2014 - The validation water extent and classification maps will be derived from high resolution optical imagery, field work, EDEN, and possibly other data. The sites for each will depend on availability of data and stratification sampling determined during pre-Validation activities.	NASA DAACs and USGS ScienceBase (ScienceBase Catalog Home).

L3 DIST	- one year time series of disturbance/no-disturbance for HLS pixels for sites across 5 strata (50 sites from each strata). - DIST_ALERT and DIST_ANN products generated for HLS products available during validation - Standard classification error analysis from accuracy assessment of DIST_ALERT and DIST_ANN (confusion matrix, accuracies) - vegetation cover grids (3 by 3 HLS pixels) generated for at least 30 sites within the US with concurrent HLS satellite overpass - Standard error analysis of vegetation cover indicator layers of DIST_ALERT including RMSE	- time-series for HLS pixels are manually labeled and identified using Planetscope data - The time stamped vegetation cover grids are derived from field work	NASA DAACs and TBD
---------	--	--	--------------------

Table 4.4: Summary of Validation deliverables for each OPERA product.

4.8.2 Validation reports and peer review

For each validated OPERA product, a final report will be generated describing the validation results at each Validation site. Prior to the final report, interim validation reports will describe progress towards this final report. The final Validation reports for each product will be peer reviewed, and made available to the science community, along with information about access to the validation data sets.

As the project proceeds, additional data products from additional sensors may be introduced into product categories. Validation reports on each new introduced product will be generated and peer reviewed as well. Product compliance with requirements will be described and compared with previous and other current products in the same product category.

To ensure that product quality continues to meet requirements during the lifetime of the project, validation reports for each product will also be generated each validation period with updated validation results that demonstrates whether continued compliance with requirements is being achieved.

References

- Agnew, D. C. (2012). SPOTL: Some programs for ocean-tide loading.
- Arundel, S. T., et al. "Preparing The National Map for the 3D Elevation Program—products, process and research." *Cartography and Geographic Information Science* 42.sup1 (2015): 40-53.
- Baetens, Louis, Camille Desjardins, and Olivier Hagolle. "Validation of copernicus Sentinel-2 cloud masks obtained from MAJA, Sen2Cor, and FMask processors using reference cloud masks generated with a supervised active learning procedure." *Remote Sensing* 11.4 (2019): 433.
- Bartos, Matt, itati, Rick Debboiut, and David Huard. 2020. Mdbartos/Pysheds: 0.2.7.1. Zenodo. <https://doi.org/10.5281/ZENODO.3822494>.
- Bekaert, David, Jermey Mauer, et al. "RAiDER: Raytracing Atmospheric Delay Estimation for RADAR". <https://github.com/dbekaert/RAiDER>. 2018--present [Accessed: Wednesday, August 4, 2021]
- Bjerklie, D.M., C.M. Birkett, J.A. Rover, J.W. Jones, C. Carabajal, J.W. Fulton, PA Garambois, J. Conway and J.T. Minear. 2018. "Satellite Remote Sensing Determination of River Discharge: Application to the Yukon River, Alaska." *Journal of Hydrology*. <https://doi.org/10.1016/j.jhydrol.2018.04.005>
- Chamberlain, Ryan, and Jennifer Schommer. "Using Docker to support reproducible research." DOI: <https://doi.org/10.6084/m9.figshare.1101910> (2014): 44.
- Claverie, M., Ju, J., Masek, J. G., Dungan, J. L., Vermote, E. F., Roger, J.-C., Skakun, S. V., & Justice, C. (2018). The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment*, 219, 145-161. [link][dataset]
- Doin, M. P., Lasserre, C., Peltzer, G., Cavalié, O., & Doubre, C. (2009). Corrections of stratified tropospheric delays in SAR interferometry: Validation with global atmospheric models. *Journal of Applied Geophysics*, 69(1), 35-50.
- Donchyts, Gennadiy, Hessel Winsemius, Jaap Schellekens, Tyler Erickson, Hongkai Gao, Hubert Savenije, and Nick van de Giesen. 2016. "Global 30m Height above the Nearest Drainage." In . Unpublished. <https://doi.org/10.13140/RG.2.1.3956.8880>.
- European Space Agency, Sinergise (2021). Copernicus Global Digital Elevation Model. Distributed by OpenTopography. <https://doi.org/10.5069/G9028PQB> Accessed: 2021-07-29
- Fattahi, H., Agram, P., & Simons, M. (2016). A network-based enhanced spectral diversity approach for TOPS time-series analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2), 777-786.
- Fattahi, H., Simons, M., & Agram, P. (2017). InSAR time-series estimation of the ionospheric phase delay: An extension of the split range-spectrum technique. *IEEE Transactions on Geoscience and Remote Sensing*, 55(10), 5984-5996.
- Werner, Charles, et al. "Gamma SAR and interferometric processing software." *Proceedings of the ers-envisat symposium, Gothenburg, Sweden*. Vol. 1620. 2000.
- Foga, Steve, et al. "Cloud detection algorithm comparison and validation for operational Landsat data products." *Remote sensing of environment* 194 (2017): 379-390.
- García-Haro, Francisco Javier, Hongliang Fang, and Juanma Lopez Sanchez. "Remote Sensing of Leaf Area Index (LAI) and Other Vegetation Parameters." (2019).
- Garthwaite, M. C., Lawrie, S., Dawson, J., & Thankappan, M. (2015). Corner reflectors as the tie between InSAR and GNSS measurements: Case study of resource extraction in Australia. *Eur. Sp. Agency. Spec. Publ., ESA SP SP-731*.
- Gisinger, C. et al., "In-Depth Verification of Sentinel-1 and TerraSAR-X Geolocation Accuracy Using the Australian Corner Reflector Array," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 2, pp. 1154-1181, Feb. 2021, doi: 10.1109/TGRS.2019.2961248.
- Guizar-Sicairos, Manuel, Samuel T. Thurman, and James R. Fienup. "Efficient subpixel image registration algorithms." *Optics letters* 33.2 (2008): 156-158.
- Hansen M.C., Egorov A., Potapov P. V., Stehman S. V., Tyukavina A., Turubanova S. A., Roy D. P., Goetz S. J., Loveland T. R., Ju J., Kommareddy A., Forsythe C., Bents T. (2014) Monitoring

- conterminous United States (CONUS) land cover change with Web-Enabled Landsat Data (WELD), *Remote Sensing of Environment*, 140, 466-484.
- Hansen M.C., Potapov P. V., Moore R., Hancher M., Turubanova S. A., Tyukavina A., Thau D., Stehman S.V., Goetz S.J., Loveland T.R., Kommareddy A., Egorov A., Chini L., Justice C.O., Townshend J.R.G. (2013) High-resolution global maps of 21-st-century forest cover change. *Science*, 342, 850-853.
- Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau et al. "High-Resolution Global Maps." (2021).
- Hansen, M.C., Krylov, A., Tyukavina, A., Potapov, P.V., Turubanova, S., Zutta, B., Suspense, I., Margono, B., Stolle, F., Moore, R. (2016) Humid tropical forest disturbance alerts using Landsat data. *Environmental Research Letters*, 11, 034008.
- Hanssen, Ramon F. Radar interferometry: data interpretation and error analysis. Vol. 2. Springer Science & Business Media, 2001.
- Hawkins, R., E. Attema, R. Crapolicchio, P. Lecomte, Josep Closa, P. J. Meadows, and S. K. Srivastava. "Stability of Amazon Backscatter at C-Band: Spaceborne Results from ERS-1/2 and RADARSAT-1." In SAR workshop: CEOS Committee on Earth Observation Satellites, vol. 450, p. 99. 2000.
- Jensen, D.J., M. Simard, R. Twilley, E. Castaneda, and A. McCall. 2021. Pre-Delta-X: Aboveground Biomass and Vegetation Maps, Wax Lake Delta, LA, USA, 2016. ORNL DAAC, Oak Ridge, Tennessee, USA. <https://doi.org/10.3334/ORNLDAAAC/1821>
- Jolivet, R., Grandin, R., Lasserre, C., Doin, M. P., & Peltzer, G. (2011). Systematic InSAR tropospheric phase delay corrections from global meteorological reanalysis data. *Geophysical Research Letters*, 38(17).
- Jones, J.W. 2011. Remote sensing of vegetation pattern and condition to monitor changes in Everglades biogeochemistry. *Critical Reviews in Environment and Technology* special edition. p Vol. 41 (S1) 64-91.
- Jones, J.W. 2015. Efficient wetland surface water detection and monitoring via Landsat: Comparison with in situ data from the Everglades Depth Estimation Network. *Remote Sensing*. Vol 7, issue 9, 12503-12538. <http://dx.doi.org/10.3390/rs70912503>.
- Jones, J.W. 2019. Improved Automated Detection of Subpixel-Scale Inundation—Revised Dynamic Surface Water Extent (DSWE) Partial Surface Water Tests. *Remote Sensing*, 11, 374. <https://doi.org/10.3390/rs11040374>
- Jones, J.W. 2020. Leveraging the Global Fiducials Library to enhance Dynamic Surface Water Extent (DSWE) Product Evaluation. AGU Special Session: The Global Fiducial Program Imagery: A Cold War legacy for today's Science. AGU Annual Meeting.
- Jones, J.W., Desmond, G.B., Henkle, C., Glover, R. 2012. An approach to regional wetland digital elevation model development using a differential global positioning system and a custom-built helicopter-based surveying system. *International Journal of Remote Sensing*, 33 (2), pp. 450-465.
- Jones, J.W., Starbuck, M.J., Shermeyer, J., Roig-Silva, C. 2016. The U.S. Geological Survey Dynamic Surface Water Extent (DSWE) Evaluation Strategy. European Union General Assembly.
- Kellndorfer, J., Cartus, O., Lavallo, M. *et al.* Global seasonal Sentinel-1 interferometric coherence and backscatter data set. *Sci Data* 9, 73 (2022). <https://doi.org/10.1038/s41597-022-01189-6>
- Landsat Level-3 Dynamic Surface Water Extent (DSWE) Science Product Digital Object Identification (DOI): /10.5066/F7445KQK
- Lehner, Bernhard, and Günther Grill. 2013. "Global River Hydrography and Network Routing: Baseline Data and New Approaches to Study the World's Large River Systems." *Hydrological Processes* 27 (15): 2171–86.
- Leister-Taylor, Vera, Philipp Jacob, Henning Schrader, and Hanjo Kahabka. 2020. "GEO1988-CopernicusDEM-RP-001_ValidationReport_I3.0.Pdf." AO/1-9422/18/I-LG. ESA.
- Li, Jian, and David P. Roy 2017. "A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring" *Remote Sensing* 9, no. 9: 902. <https://doi.org/10.3390/rs9090902>

- Liu, Yan, David Tarboton, and David Maidment. 2020. "Height Above Nearest Drainage (HAND) and Hydraulic Property Table for CONUS - Version 0.21 (20200601)." Oak Ridge National Laboratory (ORNL), Oak Ridge, TN (United States). Oak Ridge Leadership Computing Facility (OLCF); Oak Ridge National Laboratory (ORNL), Oak Ridge, TN (United States). <https://doi.org/10.13139/ORNLNCCS/1630903>.
- Luscombe, Anthony P. "Radiometric calibration information from RADARSAT-1 Amazon measurements." In *Proceedings of the 2001 CEOS SAR Workshop*, Tokyo, Japan, pp. 2-5. 2001.
- Mahdianpari, Masoud, et al. "Random forest wetland classification using ALOS-2 L-band, RADARSAT-2 C-band, and TerraSAR-X imagery." *ISPRS Journal of Photogrammetry and Remote Sensing* 130 (2017): 13-31.
- Shimada, Marc, O. Isoguchi, T. Tadono, and K. Isono, "PALSAR radiometric and geometric calibration," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 12, pp. 3915–3932, Dec. 2009.
- Simard M., B. V. Riel, M. Denbina and S. Hensley, "Radiometric Correction of Airborne Radar Images Over Forested Terrain With Topography," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4488-4500, Aug. 2016, doi: 10.1109/TGRS.2016.2543142.
- McGibbney, Lewis John, and Edward M. Armstrong. "An Innovative Open Data-driven Approach for Improved Interpretation of Coverage Data at NASA JPL's PO. DAA." *AGU Fall Meeting Abstracts*. Vol. 2016. 2016.
- Meyer, Franz J., Jeremy B. Nicoll, and Anthony P. Doulgeris. "Correction and characterization of radio frequency interference signatures in L-band synthetic aperture radar data." *IEEE Transactions on Geoscience and Remote Sensing* 51, no. 10 (2013): 4961-4972.
- Monti-Guarnieri, Andrea, Davide Giudici, and Andrea Recchia. "Identification of C-band radio frequency interferences from sentinel-1 data." *Remote Sensing* 9, no. 11 (2017): 1183.
- Morrow, Rosemary, et al. "Global observations of fine-scale ocean surface topography with the Surface Water and Ocean Topography (SWOT) mission." *Frontiers in Marine Science* 6 (2019): 232. [link]
- Nobre, A. D., L. A. Cuartas, M. Hodnett, C. D. Rennó, and S. Saleska. 2011. "Height Above the Nearest Drainage - a Hydrologically Relevant New Terrain Model." *Journal of Hydrology* 404 (1–2): 13–29.
- Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., and Wulder, M.A., 2014, Good practices for estimating area and assessing accuracy of land change, *Remote Sensing Environment*, 148.
- P. A. Rosen, S. Hensley, and C. Le, "Observations and mitigation of RFI in ALOS PALSAR SAR data: Implications for the DESDynI mission," in *Proc. IEEE Radar Conf.*, May 2008, pp. 1–6.
- Pekel, Jean-François, et al. "High-resolution mapping of global surface water and its long-term changes." *Nature* 540.7633 (2016): 418-422.
- Piantanida, Riccardo, G. Hajduch, and J. Poullaouec. "Sentinel-1 Level 1 Detailed Algorithm Definition." ESA, techreport SEN-TN-52-7445 (2016).
- Pickens, A.H., Hansen, M.C., Hancher, M., Stehman, S.V., Tyukavina, A., Potapov, P., Marroquin, B., Sherani, Z. 2020. Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series. *Remote Sensing of Environment*, 243, <https://doi.org/10.1016/j.rse.2020.111792>
- Plato, Edward A. "Highlighting Recent Uses of the NASA Worldview Mapping Application." *AGU Fall Meeting*. 2019.
- Potapov, Peter, et al. "Landsat analysis ready data for global land cover and land cover change mapping." *Remote Sensing* 12.3 (2020): 426.
- Rahmati, Omid, Aiding Kornejady, Mahmood Samadi, Antonio Donato Nobre, and Assefa M. Melesse. 2018. "Development of an Automated GIS Tool for Reproducing the HAND Terrain Model." *Environmental Modelling & Software* 102 (April): 1–12.
- Rosen, Paul, et al. "An update on the NASA-ISRO dual-frequency DBF SAR (NISAR) mission." 2016 *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 2016. [link]

- Rosen, Paul A., et al. "The InSAR scientific computing environment 3.0: a flexible framework for NISAR operational and user-led science processing." IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2018.
- Roy, David P., et al. "Landsat-8: Science and product vision for terrestrial global change research." Remote sensing of Environment 145 (2014): 154-172.
- Schmidt, Kersten, et al. "Radiometric comparison within the Sentinel-1 SAR constellation over a wide backscatter range." Remote Sensing 12.5 (2020): 854.
- Schubert, Adrian, et al. "Sentinel-1A/B combined product geolocation accuracy." Remote sensing 9.6 (2017): 607.
- Shiroma, Gustavo HX, et al. "An efficient area-based algorithm for SAR radiometric terrain correction and MAP projection." IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2020.
- Sleeter, R., Carter, E., Jones, J.W., Eggleston, J., Kroeker, S., Ganuza, J., Dobbs, K., Coltin, B., McMichael, S., Shastry, A., Longhenry, R., Ellis, B., Jiang, Z., Phillips, J., and Furlong, P. M., 2020, Satellite-Derived Training Data for Automated Flood Detection in the Continental U.S.: U.S. Geological Survey data release, <https://doi.org/10.5066/P9C7HYRV>.
- Small, David. "Flattening gamma: Radiometric terrain correction for SAR imagery." IEEE Transactions on Geoscience and Remote Sensing 49.8 (2011): 3081-3093. [link]
- Small, David. "Flattening gamma: Radiometric terrain correction for SAR imagery." IEEE Transactions on Geoscience and Remote Sensing 49, no. 8 (2011): 3081-3093.
- Smith, Deborah K., Stephanie Mullins Wingo, and Rahul Ramachandran. "Improved Open Data Sharing of NASA Airborne and Field Investigation Data." AGU Fall Meeting Abstracts. Vol. 2020. 2020.
- Torres, Ramon, Paul Snoeij, Dirk Geudtner, David Bibby, Malcolm Davidson, Evert Attema, Pierre Potin et al. "GMES Sentinel-1 mission." Remote Sensing of Environment 120 (2012): 9-24.
- Truckenbrodt, John, Terri Freemantle, Chris Williams, Tom Jones, David Small, Clémence Dubois, Christian Thiel, Cristian Rossi, Asimina Syriou, and Gregory Giuliani. "Towards Sentinel-1 SAR analysis-ready data: A best practices assessment on preparing backscatter data for the cube." Data 4, no. 3 (2019): 93. [link]
- Truckenbrodt, John, Terri Freemantle, Chris Williams, Tom Jones, David Small, Clémence Dubois, Christian Thiel, Cristian Rossi, Asimina Syriou, and Gregory Giuliani. "Towards Sentinel-1 SAR analysis-ready data: A best practices assessment on preparing backscatter data for the cube." Data 4, no. 3 (2019): 93.
- Ulaby, Fawwaz. T., Moore, Richard K., Fung, Adrian K., "Microwave Remote Sensing; Active and Passive", vol. 3, Artech House, 1986 Drusch, Matthias, et al. "Sentinel-2: ESA's optical high-resolution mission for GMES operational services." Remote sensing of Environment 120 (2012): 25-36. [link]
- Yagüe-Martínez, Néstor, et al. "Interferometric processing of Sentinel-1 TOPS data." IEEE Transactions on Geoscience and Remote Sensing 54.4 (2016): 2220-2234. [link]
- Ying, Q., Hansen, M.C., Potapov, P.V., Tyukavina, A., Wang, L., Stehman, S.V., Moore, R., Hancher, M. (2017) Global bare ground gain from 2000 to 2012 using Landsat imagery. Remote Sensing of Environment, 194, 161-176.
- Yunjun, Z., Fattahi, H., Pi, X., Rosen, P., Simons, M., Agram, P., & Aoki, Y. (2022). Range Geolocation Accuracy of C-/L-Band SAR and its Implications for Operational Stack Coregistration. IEEE Transactions on Geoscience and Remote Sensing, 60, 1-19.
- Weiss, Jonathan R., Richard J. Walters, Yu Morishita, Tim J. Wright, Milan Lazecky, Hua Wang, Ekbal Hussain et al. "High-resolution surface velocities and strain for Anatolia from Sentinel-1 InSAR and GNSS data." Geophysical Research Letters 47, no. 17 (2020): e2020GL087376.
- Werner, Charles, et al. "Gamma SAR and interferometric processing software." Proceedings of the ers-envisat symposium, Gothenburg, Sweden. Vol. 1620. 2000.

Acronyms

Acronym	Meaning
API	Application Protocol Interface
ADT	Algorithm Development Team
ASF	Alaska Satellite Facility
ATBD	Algorithm Theoretical Basis Document
AWS	Amazon Web Services
BFPQ	Block adaptive Floating-Point Quantization
CDR	Critical Design Review
cGPS	Continuous Global Positioning System
cGNSS	Continuous Global Navigation Satellite System
CSLC	Co-registered Single Look Complex (Product)
DAAC	Distributed Active Archive Center
DEM	Digital Elevation Model
DISP	Displacement product
DIST	Disturbance Product
DN	Digital Number
ECMWF	European Centre for Medium-Range Weather Forecasts
EPSG	European Petroleum Standards Group
GCP	Ground Control Point
GDS	Ground Data System
GIS	Geographic Information System
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HAND	Height above Nearest Drainage
HLS	Harmonized Landsat and Sentinel-2
InSAR	Interferometric Synthetic Aperture Radar

Acronym	Meaning
API	Application Protocol Interface
ADT	Algorithm Development Team
ASF	Alaska Satellite Facility
ISCE	The InSAR Scientific Computing Environment
L2	Level-2
L_n	Level- n , where n is an integer
LOS	Line-of-Sight
LULC	Land Use and Land Classification
LUT	Lookup Table
MGRS	Military Grid Reference System
MOE	Medium-fidelity Orbit Ephemeris
NAVSTAR	Network of U.S. satellites that provide GPS services
NISAR or NI	NASA-ISRO SAR (Mission)
PDR	Preliminary Design Review
POE	Precise Orbit Ephemeris
PRF	Pulse Repetition Frequency
PST	Project Science Team
QQP	Quasi-Quad Polarization
REMA	Reference Elevation Model for Antarctica
RFI	Radio Frequency Interference
RTC	Radiometric Terrain Correction
RTK	Real Time Kinematic Positioning
SAR	Synthetic Aperture Radar
SDS	Science Data System
SDT	Science Definition Team
SLC	Single Look Complex
SWE	Surface Water Extent (Product)
SWOT	Surface Water and Ocean Topography (Mission)

Acronym	Meaning
API	Application Protocol Interface
ADT	Algorithm Development Team
ASF	Alaska Satellite Facility
UNAVCO	University NAVSTAR Consortium

Appendix A: Sensors

There are several Earth observing sensors whose imagery is used for the OPERA data products. We briefly summarize each sensor as a background for this document. We use both SAR and optical satellite imagery. The SAR sensors used are Sentinel-1 A/B, NISAR, and SWOT. The optical sensors used are Landsat-8 and Sentinel-2 A/B. The optical data that are used for OPERA products are accessed via the [Harmonized Landsat and Sentinel-2](#) (HLS), which provide a common tiling of these distinct mission's imagery so that the data can be analyzed together [[Claverie, 2018](#)].

A.1 Sentinel-1 A/B

Sentinel-1A and -1B are a constellation of C-band SAR satellites, operated by the European Space Agency as part of the Copernicus program from the European Commission, were launched on April 3rd, 2014 and April 25th, 2016, respectively [[Yagüe-Martínez, et al. 2016](#)]. The EU commission is planning a continuation of the Sentinel-1 constellation for the next two decades with the upcoming launch of Sentinel-1 C/D both of which are scheduled for launch in 2022. OPERA will use data from Sentinel-1 A and B satellites only. Sentinel-1's Interferometric Wide (IW) swath mode has 5 meter across track by 15 meter along track resolution. IW has nearly global land coverage with a 12 day repeat per satellite and 6 day repeat pass when operated as a constellation. The IW mode is frequently used by international users and domestic services for monitoring ground displacement.

A.2 NISAR

NASA-ISRO SAR (NISAR) is a dual frequency (S- and L- band) SAR satellite that is a collaboration between NASA and ISRO; NASA is responsible for the production of the L-band sensor and ISRO, the S-band sensor [[Rosen, et al. 2016](#)]. NISAR's planned launch date is January 29th, 2023. There will be a 12-day repeat cycle capturing 5 meter resolution L-band imagery of nearly global land coverage. The relatively long wavelength (~24 meter) penetrates vegetation for broad science and monitoring applications.

A.3 SWOT

Surface Water and Ocean Topography (SWOT) Mission is a NASA mission planned for launch November 15th, 2022 and whose primary objective is to capture surface water heights globally at 50 meter resolution and 100 meters over the ocean [[Morrow, et al. 2019](#)]. SWOT will have a repeat pass frequency of about 21 days and therefore will be able to map changes of water surface over time. SWOT uses across track interferometry to measure the water surface height and will additionally capture along track heights with a nadir altimeter.

A.4 Landsat-8

The Landsat program represents the earliest and longest running earth observing satellite mission in NASA history [Roy, et al. 2014]. The current optical satellite Landsat-8 was launched February 11th, 2013 and Landsat-9 is planned to be launched September 9th, 2021 to ensure continuous temporal coverage. The satellites have near global land coverage. The satellites use the Operational Land Imager which provides 9 bands representing the visible spectrum (3 bands; RGB), near infrared (NIR), infrared (IR), short Infrared (2 bands; SWIR), panchromatic, and cirrus bands so that the RGB, NIR, IR, and SWIR are all 30 meters.

A.5 Sentinel-2 A/B

The Sentinel-2 A and B satellites are a constellation of optical satellites that were launched June 23rd, 2015 and March 17th, 2017, respectively, which together (i.e. in constellation mode) provide a 5 day repeat pass of near global land areas [Drusch, et al. 2012]. Each satellite is equipped with the Multi-Spectral Imager (MSI) with 13 bands including visible spectrum (RGB), near infrared (NIR), infrared (IR), and short Infrared (SWIR); the RGB, NIR, IR, and SWIR bands are all at 10 meter resolution.

Appendix B: Water Classes for the DSWx Product

Collectively, the Dynamic Surface Water Extent (DSWx) product will map several water related classes over a scene. The requirements for each product are in relation to these classes. The water classes with respect to a given sensor are:

1. *Open water* - an area⁷ that is entirely water and that is unobstructed to the sensor including obstructions by vegetation, terrain, or buildings.
2. *Partial surface water* - an area that is at least 50% and less than 100% open water. This may be referred to as “subpixel inundation” when referring to a pixel’s area. Examples may include inundated sinkholes, floating vegetation, or a coastal area.
3. *Inundated vegetation* - an area that contains vegetation in addition to open water, partial surface water, or both.
4. *No water* - an area that is **not** 1, 2, or 3

Note that *inundated vegetation* and *partial surface water* classes are **not** disjoint. For example, an area with sparse floating vegetation in a water body, from the definitions above, should be classified as both *partial surface water* and *inundated vegetation*. Note there are areas that are distinct to these inundated vegetation and partial surface water as well. A beach area containing only water and sand is strictly *partial surface water*, while a water body occluded entirely by vegetation is strictly *inundated vegetation*. The DSWx suite will not be responsible for differentiating these overlapping classes and each product will be responsible for a subset excluding one of these overlapping classes based on the input dataset sensitivities. Specifically, the water classes each DSWx product will identify are:

- DSWx-HLS classes:
 - *Open Water*
 - *Partial surface water*
 - *No Water*
- DSWx SAR suite (DSWx-S1, DSWx-NI, DSWx-SWOT) classes:

⁷ An area may be the area of a pixel in an image.

- *Open Water*
- *Inundated Vegetation (to varying degrees depending on sensor wavelength)*
- *No Water*

The ADT will perform a trade study to determine the feasibility and performance baseline of the DSWx-S1 and DSWx-SWOT to detect inundated vegetation. This will be discussed in future versions of this document.