# Observational Products for End-users from Remote sensing Analysis (OPERA) project

## OPERA Algorithm Theoretical Basis Document for Dynamic Surface Water Extent from Sentinel-1 A/B and NISAR data

Version 0.2.1

JPL D-108763 Rev A

May 26, 2024

Authors: Jungkyo Jung, Heresh Fattahi

Jet Propulsion Laboratory
California Institute of Technology

# SIGNATURE PAGE

## Prepared by:

*Email approval on file in EPDM*

Jungkyo Jung
OPERA Algorithm Development Team
DSWx-S1 Cognizant Engineer

Date

*Electronic Approval in EPDM*

Heresh Fattahi
OPERA Algorithm Development Team Lead

Date

## Approved by:

*Electronic Approval in EPDM*

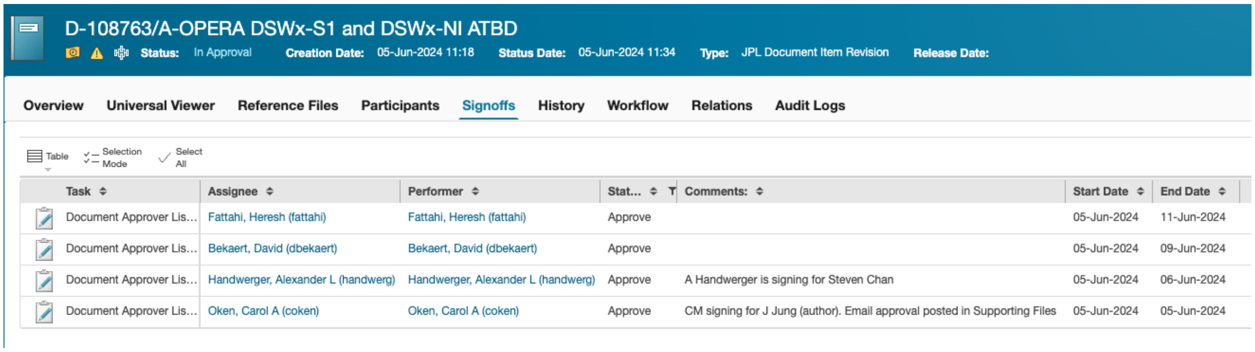
David Bekaert
OPERA Project Manager

Date

*Alexander Hardwenger Electronic Approval in EPDM for S. Chan*

Steven K. Chan
OPERA Project Scientist

Date



D-108763/A-OPERA DSWx-S1 and DSWx-NI ATBD

**Status:** In Approval **Creation Date:** 05-Jun-2024 11:18 **Status Date:** 05-Jun-2024 11:34 **Type:** JPL Document Item Revision **Release Date:**

| Task | Assignee | Performer | Stat... | Comments: | Start Date | End Date |
|---|---|---|---|---|---|---|
| Document Approver Lis… | Fattahi, Heresh (fattahi) | Fattahi, Heresh (fattahi) | Approve | | 05-Jun-2024 | 11-Jun-2024 |
| Document Approver Lis… | Bekaert, David (dbekaert) | Bekaert, David (dbekaert) | Approve | | 05-Jun-2024 | 09-Jun-2024 |
| Document Approver Lis… | Handwerger, Alexander L (handwerg) | Handwerger, Alexander L (handwerg) | Approve | A Handwerger is signing for Steven Chan | 05-Jun-2024 | 06-Jun-2024 |
| Document Approver Lis… | Oken, Carol A (coken) | Oken, Carol A (coken) | Approve | CM signing for J Jung (author). Email approval posted in Supporting Files | 05-Jun-2024 | 05-Jun-2024 |

# DOCUMENT CHANGE LOG

| Revision | Cover Date | Sections Changed | ECR # | Reason, ECR Title, LRS #* |
|---|---|---|---|---|
| Initial revision | June 15, 2023 | All | N/A | New document, LRS# LRR073169 |
| V0.2 | May 26, 2024 | All | N/A | Change title and add the DSWx-NI |
| V0.2.1 – Rev A | May 26, 2024 | Cleared document for unlimited release | N/A | CL#24-2842 |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

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

# TBD / TBC LOG

| Section | Description | Due Date |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

# TABLE OF CONTENTS

**1 Introduction** **5**
**2 DSWx-S1 and DSWx-NI Algorithm Description** **6**
  2.1 Background 6
  2.2 Algorithm description 7
    2.2.1 Mosaic 7
    2.2.2 Initial threshold computation 8
      2.2.2.1 Image tiling 8
      - 2.2.2.2. Tile selection 9
      2.2.2.3. Threshold computation 11
    2.2.3 Fuzzy logic computation 13
    2.2.4 Region growing 14
    2.2.5 False positive removal using external data 14
    2.2.6 False positive removal using bimodality test 15
    2.2.7 Inundated vegetation delineation 16
      2.2.7.1. Inundated vegetation mapping from C-band Sentinel-1 data 16
      2.2.7.2. Inundated vegetation mapping from L-band NISAR 17
  2.3 Algorithm Inputs 17
  2.4 Algorithm Outputs 18
**3 Algorithm Assumptions** **19**
  - Static ancillary data 19
**4 Algorithm Implementation** **19**
**5 Algorithm Usage Constraints** **19**
**6 Performance Assessment Validation** **20**
  6.1 Performance Assessment Validation Methods 20
  6.2 Performance Assessment Validation Uncertainties 22
  6.3 Performance Assessment Validation Errors 22
**7 Data Access** **22**
**8 Contacts** **23**
**9 References** **23**
**10 ACRONYMS** **28**
**11 List of symbols** **29**
**12 Acknowledgements** **29**

# 1 INTRODUCTION

Monitoring water extent and their temporal variations are essential information to understand local and global hydrological processes and enable efficient management of water resources. The water extent maps can provide an important role in the calibration and validation of hydraulic models for river flow processes (García-Pintado et al., 2015; Matgen et al., 2011). Furthermore, water extent maps can serve as crucial tools for effective disaster management. They assist decision-makers in responding to extreme, abnormal fluctuations in water resources, such as floods, that pose a threat to human lives and have significant potential for causing economic damage. (Merz et al., 2010; Messner and Meyer, 2006).

The need for reliable and detailed information about the spatiotemporal changes in the water resource has motivated the development of methodologies and sensors. The ground-based gauges provide the hourly or daily status of the water elevation and enable monitoring of real-time water levels with sufficient accuracy along the water reservoirs and rivers (Lesack and Melack, 1995). However, the observations are constrained to water elevations and do not provide an overview of water extent. Furthermore, these data are difficult to assess where in-situ monitoring networks are insufficient (Alsdorf et al., 2007). In this perspective, the satellite remote sensing data play a complementary role by enabling estimation of the water extents with cost-effective advantage.

Surface water extent estimation methods have been implemented using optical sensors such as Landsat, Sentinel-2, and the Moderate Resolution Imaging Spectroradiometer (MODIS) for decades. These sensors rely on spectral indices derived from multispectral data to distinguish water features from surrounding areas (Brakenridge and Anderson, 2006; DeVries et al., 2017; Du et al., 2014; Jones, 2019; Khandelwal et al., 2017; Wieland and Martinis, 2020). Optical sensors have shown promise in mapping water boundaries with sufficient accuracy. However, water extent estimation from optical data is limited to daylight and cloud-free acquisitions. In the case of flooding circumstances, cloud cover often prevents the delineation of the inundated area, despite the high demand for water extents. This limitation may also cause data gaps in time-series measurements, making it difficult to understand the seasonality of water bodies.

Synthetic Aperture Radar (SAR) sensors offer significant advantages over optical sensors for water extent monitoring. SAR sensors can illuminate the surface regardless of cloud coverage, in all weather conditions and during day and night, enabling higher temporal resolution observation of surface water. As a result, numerous methodologies have been proposed to incorporate SAR images to monitor temporal changes in water extent, including flood detection applications (Chini et al., 2017; Clement et al., 2018; Elkhrachy, 2018; Hostache et al., 2018; Psomiadis, 2016; Pulvirenti et al., 2013; Tay et al., 2020). The Sentinel-1 constellation has collected a rich archive of SAR images at C-band with 6-days or 12-days intervals, covering most land masses on the globe. Several studies have explored the potential of Sentinel-1 C-band data to delineate water bodies, with highly promising results (Bauer-Marschallinger et al., 2021; Kellndorfer et al., 2022; Li et al., 2020). The basic idea behind using SAR images to delineate water bodies is that open water surfaces have strong specular reflections, resulting in very weak received signals compared to other non-water targets.

The Dynamic Surface Water Extent from Sentinel-1 (DSWx-S1) and NISAR (DSWx-NI) algorithm generates geographic data that indicate whether surface water is present at the time of input image collection with a spatial resolution of single image pixels. This document describes the DSWx-S1 and DSWx-NI algorithm as implemented for near-global processing by the NASA-funded JPL project named Operational Products for End-users from Remote Sensing Analysis (OPERA). Following treatments of algorithm lineage and scientific underpinning, the required processing is efficiently described in a manner that facilitates implementation.

# 2 DSWX-S1 AND DSWX-NI ALGORITHM DESCRIPTION

## 2.1 Background

Many methods and examples for mapping water bodies, in particular the flooded area, have been introduced. Some of their techniques are motivated by the change detection algorithm, which compares the backscattering level difference between the pre-flood and co-flood status (Giustarini et al., 2012). To highlight the flooded areas, some research rely on the thresholding methods after computing the log ratio between pre-flood and co-flood images (Chini et al., 2017; Tay et al., 2020).  The multi-temporal analysis can be employed to detect the abnormal change in the amplitude due to the flood status compared to the statistics of the pre-flood status (Clement et al., 2018; Lin et al., 2019; Uddin et al., 2019). However, one of the limitations of these change detection-based water detection algorithms mainly aims to detect the changes of water extent compared to the reference images. Thus, the reference water boundary should be accurately determined to map the standing water. In addition, the complication of the interpretation remains whether the detected changes are induced by the new water or different sources of the changes.

Unlike the change detection method, the automation of the open water detection from a single SAR image has been developed based on thresholding methods under the assumption that backscatter values of the water area in SAR images are low so that the appropriate threshold can delineate the water from non-water areas (Chini et al., 2017; Ding and Li, 2011; Liang and Liu, 2020; Liu and Jezek, 2004; Martinis et al., 2009; Nakmuenwai et al., 2017). This method has been known as a simple and intuitive way to solve problems by showing reasonable accuracy to detect water bodies.

Despite efforts to extract water bodies from SAR images, there are several remaining issues that can severely affect the accuracy of detection. One such issue is the presence of surfaces with similar backscattering characteristics to water bodies, which can lead to commission or omission errors. For example, sandy surfaces, which typically have low backscattering levels, may be misclassified as open water (Martinis et al., 2018), as can dry bare soils (Martinis et al., 2015; Ouled Sghaier et al., 2018) or smooth surfaces such as pavement. To address this problem, Martinis et al. (2018) introduced the Sand Exclusion Layer (SEL), which is derived from time-series statistics of amplitude images and can separate sandy surfaces from flooded surfaces. However, this method requires time-series images to compute the statistics of each pixel and assumes that the surface characteristics are low and unchanged in the time series. Thus, the method may not be fully understood in areas where surface characteristics exhibit temporal variations in backscattering, or where low backscattering is observed only during limited seasons.

## 2.2 Algorithm description

The DSWx-S1 and DSWx-NI algorithms generate the surface water extent products from OPERA RTC-S1 and NISAR GCOV products, respectively. These inputs differ in center frequencies, polarizations, and pixel spacing, necessitating distinct functionalities to process and create consistent products in DSWx-S1 and DSWx-NI. Additionally, due to the differences in penetration depth between the L-band used by NISAR and the C-band used by Sentinel-1, the mapping of inundated vegetation may vary. However, the workflows to delineate open water from the RTC-S1 and GCOV products share common functionalities. The DSWx-S1 and DSWx-NI algorithm can be summarized into the following steps:

1. *Mosaic*:
    a. DSWx-S1: Merge all RTC burst products to generate a mosaic rtc image over the area of interest which usually contains multiple MGRS tiles.
    b. DSWx-NI: Merge all GCOV products to generate a mosaic image over the area of interest.
2. *Initial threshold computation*: Split the entire image into the tiles, check if the tile has the boundary between water and non-water bodies, and compute the spatially varying thresholds.
3. *Fuzzy logic computation:* Compute the fuzzy logic values utilizing spatially varying local thresholds, HAND, terrain slope angle, and reference water.
4. *Region growing:* Iteratively dilate the binary objects, starting from the seed pixels until the tolerance is reached.
5. *False positive removal using ancillary data:* From the reference water image, and land cover map, extract the potential low-backscattering areas, and compare with the co- and cross-polarization image. Mask out the low-backscatter pixels which are unlikely to be the water.
6. *False positive removal using bimodality test:* Remove the misclassified water pixels using statistical bimodality behavior with the pixels along the water bodies.
7. *Inundated vegetation delineation*: Detect the inundated vegetation from co-polarization to cross-polarization ratio.
8. *package output products to MGRS tile.*

The following subsections provide a detailed description of the different steps of the algorithm.

### - 2.2.1 Mosaic

The first step of the DSWx-S1 algorithm involves mosaicking all available RTC-S1 burst products to create a comprehensive RTC image that covers the entire area of interest. This step is crucial because it ensures that the computed thresholds (2.2.2), which distinguish between water and non-water areas, are based on reliable data. The underlying assumption is that the RTC-S1 image will display clear boundaries between water and non-water bodies, with separable distributions from the two classes.

A single RTC-S1 burst product typically covers a relatively small region, approximately 20 x 80 km. This limited coverage may not provide enough pixels to accurately represent the boundary between water and non-water areas, potentially leading to failures in threshold computation. However, when mosaic images are created from multiple bursts, they tend to cover larger areas. This increased coverage enhances the likelihood of the presence of water and non-water regions in the data and therefore increases the probability of detecting water bodies and their boundaries, thereby improving the reliability of the threshold computations to estimate water bodies.

For the DSWx-NI algorithm, which utilizes NISAR GCOV products, the process involves creating mosaic RTC images from these inputs. The OPERA RTC-S1 products have a pixel spacing of 30 m, whereas NISAR GCOV products have variable pixel spacings depending on the bandwidth. Specifically, the GCOVs with 80 MHz, 40 MHz, and 20 MHz bandwidths have pixel spacings of 20 m, 10 m, and 20 m, respectively. When it is necessary to merge multiple GCOV products with diverse pixel spacings, the coarser GCOV (20 m) is oversampled to 10 m. This resampling preserves the detailed features of the water bodies, which is essential for accurate water detection.

In cases where the input RTC-S1 or NISAR GCOV products have varying polarizations, mosaicking is not required because the backscattering characteristics of water bodies differ depending on the polarization. This variance means that each polarization can provide unique and valuable information about the presence and characteristics of water bodies. In this situation we estimate water bodies from each polarization separately and mosaic the estimated water products at the end.

### - 2.2.2 Initial threshold computation

#### - 2.2.2.1 Image tiling

Estimating the surface water extent from single SAR data acquired from imaging geometries with moderate to high incidence angles, such as Sentinel-1 and NISAR, relies on the fact that the surface water bodies exhibit low backscattering due to the strong specular reflection, meanwhile, the non-water bodies have relatively higher backscattering. This property leads to darker appearance of water bodies and brighter appearance of non-water targets in such SAR images. Therefore, areas with surface water can be estimated by thresholding the SAR backscatter.

Since the Sentinel-1 images cover almost 240 km in range direction, consequently a frame of SLC image contains a variety of land cover types with diverse backscattering characteristics. Furthermore, if a small part of the image covers the water bodies but a large part covers the non-water areas, the distributions of the two classes are strongly imbalanced. If the relatively small distribution of the water bodies on land is dominated by the non-water targets, then the SAR backscatter distribution of the water pixels is not distinguishable from the distribution of land pixels in a histogram of the SAR backscatter data and the threshold estimation may result in ill-posed thresholds leading to mis-classification of water pixels (Chini et al., 2017; Gong et al., 2016). To avoid this problem, a tile-based thresholding algorithm is implemented (Chini et al., 2017; Martinis et al., 2015; Twele et al., 2016).

In the tile-based thresholding algorithm, the tiles to be used for thresholding should contain two distinct and separable distributions coming from water and non-water pixels. Therefore, the threshold can be determined where the two different distributions coexist in the histogram. The algorithm has two main steps: 1) split the entire image into smaller tiles, and 2) select the tiles showing the clear bimodality of the distribution. The first step in this process is to constrain the spatial extent by dividing the entire image into tiles, referred to as S0. Given that the backscattering returned from the surface water is influenced by the environment conditions, the appropriate threshold can vary from tile to tile. By partitioning the entire image into smaller tiles S0 , it becomes possible to calculate the spatially varying threshold. This subdivision not only accommodates the diversity of conditions across the image but also enhances the precision of the thresholding process.

In finding the tiles exhibiting the bimodality, the appropriate tile size varies due to the spatially diverse sizes of the water objects. To capture the bimodality behavior only in small tiles, the size and location of the tiles need to be adjustable. Each tile $S_0$ is split into half-sized sub-tiles $S_1$. The sub-tile $S_1$ shifts within tile $S_0$ with the steps of $S_1/2$ in a loop to find the sub-tiles satisfying the bimodality criteria (Section 2.2.2.2). The number of the sub-tiles satisfying the conditions is recorded and the procedure is terminated once enough tiles are found. If the number of tiles meeting the conditions is under the criteria, then smaller sub-tiles $S_2$ having half the size of $S_1$ are created and the procedure is repeated. The procedure is iterated until a sufficient number of tiles are obtained. Then, the thresholds can be computed from the selected tiles showing the bimodality (section 2.2.2.3.). If the splitting procedure continues up to the limit of the minimum tile sizes, and still fails to ensure enough number of tiles, then the tile $S_0$ is considered to have a unimodal distribution of SAR backscatter and will not be used to estimate water/no-water threshold.

### - 2.2.2.2. Tile selection

To identify the bimodal distribution in the sub-tiles, $S_1, S_1, S_1, \ldots, S_N$, the following measure can be used:

$$CVx = \frac{\sigma_{sub-tile}}{\mu_{sub-tile}} \quad (1)$$

The $CVx$ represents the ratio between the standard deviation and the mean value computed from pixels in the individual sub-tiles. Hence, the ratio appears high where the radar backscattering is heterogenous and appears low where the radar backscattering is homogeneous. When the $CVx$ value is higher than 0.1, these conditions are typically met along the boundary of a tile where water and non-water areas coexist. This phenomenon serves as an indicator that the tiles contain more than one semantic class (Martinis et al., 2009). However, in regions characterized by high average backscatter and high variance such as urban areas or agricultural fields, reliance on the *CVx* metric may lead to false selection of the tile.

To mitigate the risk of false tile selection, another metric is introduced, *Rx* as,

$$Rx = \mu_{sub-tiles}/\mu_{S_0} \qquad (2)$$

The lowest *Rx* values are typically found in tiles containing only water bodies, such as oceans, while higher values are observed over urban or agricultural areas. tiles with *Rx* < 0.98 are added to the candidate list.

The Sarle's bimodality coefficient is defined as (Ellison, 1987; Pfister and Janczyk, 2013).

$$b = \frac{g^2+1}{k+\frac{3(n-1)^2}{(n-2)(n-3)}} \qquad (3)$$

where n is the sample size, g is the sample skewness of the distribution and k is the sample excess kurtosis. The skewness and kurtosis are corrected for sample bias. When the Sarle's bimodality coefficient is below 5/9, the distribution is typically expected to be the uniform distribution. Higher coefficients are likely to have the bimodal distribution. However, the test would fail when the distribution is heavily skewed by leading to high value (Pfister and Janczyk, 2013).

Assuming, the bimodal distribution at a tile is a sum of two Gaussian distributions, additional bimodality tests can be adopted as

$$f(x) = A_1\, exp\left(-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right) + A_2\, exp\left(-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2\right) \qquad (4)$$

where $A_1$ and $A_2$ are amplitudes of gaussian curves, $\mu_1$ and $\mu_2$ are the mean values of the distributions, and $\sigma_1$ and $\sigma_2$ are the standard deviations. The mean and standard deviation for two Gaussian curves are estimated from the Levenberg-Marquardt algorithm. The initial values for each parameter are given from the two separated Gaussian distributions, which are divided using Otsu thresholding methods. Here, it is assumed that $\mu_1$ is always lower than $\mu_2$ and the first and second modes represent the water and non-water targets, respectively.

If the Gaussian curve fitting is successfully converged, and their parameters are obtained, then the metrics to quantify the bimodality of the distribution can be computed. Ashman D coefficient describes the separation of the means of the two Gaussian distributions relative to their width and can be expressed as (Ashman et al., 1994)

$$D = \sqrt{2}\frac{|\mu_1-\mu_2|}{\sqrt{(\sigma_1{}^2+\sigma_2{}^2)}}. \qquad (5)$$

If D is higher than 2, the distribution is assumed to have a clear bimodality.

The next metric is the surface ratio between two distributions. The area under a Gaussian distribution curve can be expressed as $A\sigma\sqrt{2\pi}$. The areas of two possible Gaussian distributions in a tile should be balanced enough to have their own statistical representativeness. In order to examine the ratio of the surface under the two Gaussian curves, the Surface Ratio (SR) parameter is defined as

$$SR = \frac{min\,(A_1\sigma_1\sqrt{2\pi}, A_2\sigma_2\sqrt{2\pi})}{max\,(A_1\sigma_1\sqrt{2\pi}, A_2\sigma_2\sqrt{2\pi})} \quad (6)$$

The Ashman D coefficient is meaningful only when SR between the two found curves is high enough. If SR is higher than 0.1, the condition is fulfilled (Chini et al., 2017). If all metrics estimated from equations (1)-(6) satisfy the condition, then the bimodality behavior can be assumed in the distributions. In the case that the Gaussian curve fitting fails, only three tests (1)-(3) are evaluated.

The tile selection strategy based on the bimodal distribution test is a standalone method that utilizes image intensity. However, the heterogeneous and diverse backscattering characteristics of non-water targets can lead to behaviors similar to those of water, such as in farmland and forests.

To minimize the error of selecting the incorrect tiles only containing non-water targets, the external information on the surface water extents can be jointly incorporated. The surface water information (e.g., SRTM water mask, or JRC global surface water occurrence or seasonality map (Pekel et al., 2016)) is freely and globally accessible and can be easily integrated into the tile selection steps. Hence, if the tiles do not cover the water and non-water boundaries, then the tiles can be ruled out from the selection step even though the bimodal behavior is observed. The water boundary is defined from the JRC surface water occurrence map by finding the occurrence value higher than 5.

-

### - 2.2.2.3. Threshold computation

After selecting the tiles that contain water and no-water pixels, then thresholds discriminating the water and non-water can be computed for individual selected sub-tiles. In most of the cases, the computed thresholds in this step delineate the boundary between water and non-water areas. However, the thresholds are estimated assuming that the distribution of the backscatter within a tile follows a bimodal distribution. This assumption may be invalid for tiles where more than two semantic classes co-exist. This situation may occur when a tile contains low-backscattering land, high-backscattering land, and water. Another example is a tile with three classes of low-backscattering water, bright backscattering water and high backscattering land. In these edge cases, the bimodal assumption leads to significant increase of false positive and false negative estimates.

If three distinguishable distributions exist in a tile, then equation (4) can be reformulated as,

$$f(x) = A_1\,exp\left(-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right) + A_2\,exp\left(-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2\right) + A_3\,exp\left(-\frac{1}{2}\left(\frac{x-\mu_3}{\sigma_3}\right)^2\right) \quad (7)$$

The Gaussian curve fitting procedure is re-carried out to find the modes from the histogram. The initial values for means are obtained from the multi-level Otsu threshold algorithm (Liao et al., 2001). Once Gaussian curve fitting converges, then, the three approximated Gaussian distributions are obtained. If all the three modes have high amplitudes (i.e., $A_1 / A_{max}$ > 0.1 AND $A_2 / A_{max}$ > 0.1 AND $A_3 / A_{max}$>0.1), it is assumed that the tri-modal distribution is observed. If not, the distribution is assumed to be bimodal. In the tri-modal distribution case, two possible scenarios arise. First, if the semantic classes are related to water, low-backscattering non-water, and high-backscattering non-water, the lower threshold will be located between water and low-backscattering non-water. Second, Second, if the classes are low-backscattering water, high-backscattering water due to windy conditions, and non-water, then the lower threshold is placed between the modes of water bodies. The working hypothesis assumes that the cross-polarization channel is less affected by wind-driven roughening effects, meaning that the latter case is not observed and all water bodies will be detected in cross-polarization. Thus, it is beneficial to choose the lower threshold in both scenarios.

In instances where one of the mode amplitudes is negligible, the distribution can be reasonably assumed to be bimodal. The Kittler and Illingworth minimum error thresholding approach can be adopted to calculate the threshold from bimodal distributions (Kittler and Illingworth, 1986). The threshold computed in this step is typically positioned between two modes fitted in equation (4). In addition to the thresholds, the peak value of the first mode is stored. This is because both values are intended for constructing the fuzzy set described later in section 2.2.3.

The thresholds and peak values computed from the individual sub-tiles within each root tile $S_0$ are averaged and those values are assigned to the root tile $S_0$ covering the sub-tiles. The thresholds of root tile $S_0$ are interpolated to spatially extend to entire images even where the water bodies are not found in the reference water map. This is because newly inundated areas due to the flood may not be recorded in the reference water map. Furthermore, the spatially varying threshold enables the preservation of the optimized thresholds for each water body on diverse land cover types.

### - 2.2.3 Fuzzy logic computation

The threshold-based approach for determining water binary often encounters difficulties in accurately classifying low-backscattering non-water areas (hereafter referred to as dark lands) and high-backscattering water areas (hereafter referred to as bright water). The implementation of a fuzzy logic-based approach can enhance classification accuracy by eliminating water-lookalike pixels and incorporating land-lookalike pixels situated above water bodies. In this study, co- and cross-polarization intensity, local slope angle, height above the nearest drainage (HAND), and existing water occurrence maps are employed as components for constructing the fuzzy set. Each element is rescaled and normalized using standard Z- and S-shape membership functions. The S- and Z-shaped membership functions for each component necessitate two parameters for foot and shoulder values, which are independently determined for each component. The thresholds determined from the previous step and their interpolated peak values are shoulder and foot for the Z-membership function for co- and cross-polarized RTC intensity. Z-membership functions are also applied to the local slope angle [0°, 15°] and HAND [0 m, 200 m]. The S-membership function is utilized for the Joint Research Centre (JRC) surface water occurrence map to ascertain that water has persisted for 37 years. The maximum grade is assigned to the pixels with a value higher than 80, and the minimum grade is to the pixel lower than 5 of occurrence value.

The composite fuzzy set is created by integrating all fuzzy elements into a single fuzzy set, achieved  by averaging all membership degrees. The value 1 in the composite fuzzy set signifies that all fuzzy components exhibit a high likelihood of water presence, while a value of 0 indicates an absence of water. For example, the objects where water existed over the past decades but no longer exists due to the drought could have the membership value of 1 for the slope, HAND, and water occurrence map, but have higher backscattering values in co- and cross-polarization RTC image. This results in a composite fuzzy set value of 0.6, leading to a classification as non-water. On other hand, the inundated areas due to the flood may lead to lower membership values for the water occurrence map but higher values for memberships of intensity, slope angle, and HAND. Thus, the composite fuzzy set value reaches to 0.8, which is sufficient for classification as water bodies.

### - 2.2.4 *Region growing*

To increase the homogeneity of the water plain and identify the water along the boundary between water and non-water, an additional region growing step is applied to the composite fuzzy set. The region-growing step is a procedure that iteratively dilates the binary objects, starting from the seed pixels until the tolerance is reached. In order to avoid the false detection of the water-lookalike object, a strict value of the composite fuzzy set of 0.8 is recommended, representing only objects which highly fulfill the conditions of the existence of water. Thus, the dilation starts from the object is likely water until the tolerance, which is the relaxed threshold, 0.6 reached.  The dilation is only applied to the nearest neighbor pixels from the previous iteration assuming that the water pixels are spatially connected to seed pixels rather than spread out over the entire image. Following the region-growing method, a binary image containing water and non-water candidates can be obtained.

### - 2.2.5 *False positive removal using external* data

One of the most challenging parts of water mapping is to distinguish the dark land areas from the true water bodies, which can potentially lead to false positives in the classification results. Even after the fuzzy-logic computation and region growing, the binary water map may suffer from dark land areas. The dark lands are typically observed over flat surfaces with small roughness, e.g., paved roads, where strong specular reflections are dominant. Also, some surface types such as sand could show low backscattering due to their material characteristic absorbing the microwave signals. Often, the sand is even darker than water bodies (Bauer-Marschallinger et al., 2021).

The comparison of the C-band backscattering characteristics against diverse land cover types revealed that the intensity distributions of water bodies are significantly overlapped with some land covers such as bare/sparse vegetation, herbaceous vegetation, and moss & lichen. These similar backscattering characteristics hamper distinguishing the water from dark lands with a thresholding approach. Since the arid areas such as subtropics in Africa, Australia, and China showing very low backscatter are mainly classified as bare/sparse vegetation, the high false rate may be observed in those regions even after the fuzzy-computation step.

Based on the analysis above, the idea utilized in the DSWx-S1 and DSWx-NI algorithm is that the dark land candidates, which are misclassified as water bodies due to their low backscattering intensity values, could be extracted from land cover maps. However, relying solely on land cover maps may result in ambiguity when mapping the dynamic surface water bodies. Given that the ESA worldcover map is created from optical imagery and the radar backscatter, the land cover layers contain not only dark land candidates, but also the high backscattering regions. Hence, dark lands candidates extracted from land cover need to be refined by combining the dual-polarization intensity information to avoid redundant masking.

The selection of an appropriate threshold to define the dark land candidates is essential, the thresholds were derived through comparison between the more than 30 RTC scenes and ESA World cover maps of different land covers and hydrological environments. The histograms of the permanent water and the other land cover except bare/sparse vegetation for intensity of RTC show that surface water can be separable under the following condition.

$$\gamma_T^0(VV) < -13.4 \text{ dB} \qquad (8)$$

$$\gamma_T^0(VH) < -22.2 \text{ dB} \qquad (9)$$

Hence, the areas with the intensity values lower than above conditions are likely to be close to either water or dark land candidates.

Even though the dark land candidates are extracted from the RTC intensity image and landcover map, inaccuracies may arise in areas where the water extent variation is high due to the flood, tide, or seasonal inundation. For example, areas that are seasonally inundated, might be classified as bare/sparse ground and herbaceous vegetation. When the areas are flooded and have very low backscattering, the dark land candidates extracted from the two pieces of information may identify them as dark land candidates. Thus, to avoid masking out the natural water bodies, the seasonality of the water extent can be jointly used. The JRC surface water seasonality product ranges from 0 to 12, representing no water and permanent water, respectively. By identifying the areas with 0, from the water seasonality map, the areas where the water unlikely exists could be refined based on the history of 37 years.

The final dark land candidates are determined by applying a logical "AND" operator to the binary layers of intensity, land cover, and water occurrence map. The underlying hypothesis is that the intersecting areas from above conditions are likely not related to the surface water bodies.

### - 2.2.6 *False positive removal using bimodality test*

The previous step corrects for dark and non-inundated areas which are classified as bare/sparse vegetation in the land cover maps. However, some dark land areas may remain after the masking based on the land cover map. The dark land mask generated from the land cover map is selected for the specific land cover type. Thus, the low backscattering areas placed on the other land covers not being used for masking may be left as the residuals. In addition, some land types exhibit seasonal variations of vegetation density, and they can become sparse/bare ground in certain seasons even though classified as vegetated areas on Worldcover map. Thus, an additional step to distinguish between the dark land and water class without solely relying on the land cover map is necessary.

The basic assumption for distinguishing between water bodies and dark lands is that areas comprising both non-water and water will exhibit a bimodal distribution, while residual dark lands and surrounding areas will show a unimodal distribution. The anticipated behaviors in the dark lands and adjacent areas are smooth intensity changes in the space domain, and insignificant intensity differences between the dark land and surrounding land. In contrast, water bodies and surrounding pixels exhibit the opposite behavior. Therefore, the presence of bimodal behavior is a key parameter for refining water bodies by eliminating dark lands exhibiting unimodality. Unlike the tile-selecting step, the procedure is carried out with remaining binary objects and surrounding pixels instead of the rectangular-shaped sub-tile window. To quantify the bimodality, the dilation is applied to binary objects by progressively expanding the boundaries of foreground pixel regions to encompass the surrounding pixels. The iterations must continue until the number of dilated pixels is balanced with the number of pixels within the binary object, ensuring statistical representativeness. The iterative dilations extend to non-water areas without the binary objects to prevent the statistical bias.

Based on the bimodality tests introduced in Section 2.2, if all coefficients estimated from (3)-(6) meet the conditions, the dark land candidates can be assumed to be the water body. Otherwise, the dark land candidates are confirmed as dark land and masked out from water class. The bimodality test needs to run only for uncertain and controversial areas. For instance, the areas with low backscattering below thresholds in co- and cross-polarizations and high-water occurrence values are likely to be water. However, if the area has low backscattering but low water occurrence, this area could be a new water body or dark land. To identify and list the candidates taken to be the bimodality test, the defuzzified image can comprehensively provide the uncertainty level. Hence, the binary objects containing the value of 1 in the defuzzified image were assumed to be spatially grown from seed in fuzzy-logic based strategy. These are potentially considered water bodies. However, the water body candidates containing only below the defuzzified value of 1, then they are enrolled to be listed in the bimodality test. The uncertainty can be clarified from the bimodality test by determining whether it is close to water or non-water.

### - 2.2.7 Inundated vegetation delineation

#### - 2.2.7.1. Inundated vegetation mapping from C-band Sentinel-1 data

The detection of water bodies under vegetation presents challenges as they often exhibit higher backscattering compared to open water bodies. Consequently, the algorithm and assumption designed for open water is not directly applicable to inundated vegetation.

Based on the fact that the inundated vegetation induces the double-bounce scattering or volume scattering while the open water generally causes specular reflection resulting in the low backscatter in the radar signal. Hence, the co-polarization over the inundated vegetation areas typically has the high backscattering. The co-polarization to cross-polarization ratio is a key estimate to classify the inundated vegetation. In particular, normalizing by VH removes the effects of topography (leftover uncompensated effects after radiometric terrain calibration) and biomass, since these affect both VV and VH similarly.

After computing the polarimetric ratio between co- and cross-polarization, the inundated vegetation is mapped by applying the threshold specific to the inundated vegetation.

High value in the polarimetric ratio is expected not only on the inundated vegetation, but also over the many areas including urban regions. To minimize the misclassification, the ancillary data is used to refine the results. Hence, the inundated vegetation is limited to the Herbaceous wetland that the ESA world cover map defines.

#### - 2.2.7.2. Inundated vegetation mapping from L-band NISAR

Similar to the inundated vegetation method in the C-band Sentinel-1, the polarimetric ratio is the main estimate to classify the inundated vegetation. However, the L-band leads to detection of the extensive inundated vegetation due to deeper penetration and interaction with internal structure. Thus, unlike that the inundated vegetation is limited to the Herbaceous wetland in C-band, the DSWx-NI includes the Forest, Herbaceous wetland, Shrubs, Grassland, and Mangrove.

The double-bounce scattering coming from the interaction with the water and vegetation could be different depending on the characteristics of the targets. So, specific thresholds for the different landcover types could be variable.

## 2.3 Algorithm Inputs

The main input and ancillary inputs to the DSWx-S1 algorithm are summarized in Table 2. It includes :

- OPERA RTC-S1 (DSWx-S1) or NISAR GCOV (DSWx-NI) product

  The algorithm receives multiple OPERA RTC products or NISAR GCOV products as inputs.

- Land Cover (WorldCover 2.0)

  The WorldCover product contains broad land cover classes with 10-m spatial resolution in global areas. The classes consist of "Tree cover", "Shrubland", "Grassland", "Cropland", "Built-up", "Bare/sparse vegetation", "Snow and ice" "Permanent water bodies", "Herbaceous wetland", "Mangroves" and "Moss and lichen". It covers global regions excluding Antarctica.

- DEM (Copernicus GLO-30)

  The algorithm uses the project-specified DEM for processing based on the Copernicus DEM 30 meters (GLO-30 DEM). The DEM is the same as prepared for the NASA-ISRO Synthetic Aperture Radar mission, referred to as the NISAR-DEM. This DEM can be overridden by an external user-defined DEM.

- HAND

  The DSWx-S1 uses the Height Above Nearest Drainage (HAND) for processing DSWx-S1 workflow and masking WTR, and BWTR layers. The HAND has a 30-m spacing. The data is prepared from ASF and publicly available through the S3 bucket.

- Reference Surface water map

  The SAS uses the JRC Global Surface Water data to process DSWx-S1. The product consists of 'occurrence', 'seasonality', 'recurrence', and 'transitions' dataset for global surface water derived from collections of optical images. The DSWx-S1 uses 'occurrence' or 'seasonality' for ancillary input.

Table 2. Inputs to the DSWx-S1 algorithm.

| Input Data | Description | Producer |
|---|---|---|
| OPERA RTC S1 | Multiple RTC-S1 products consisting of burst and metadata | JPL |
| WorldCover 2.0 | Worldcover maps containing 11 land cover classes with global coverage at 10 m resolution. | ESA |
| Copernicus GLO-30 | Digital Elevation Model with global coverage at 30 m resolution. | ESA |
| HAND | Height above nearest drainage data | ASF |
| Surface water seasonality map | Surface water reference map with values ranging from 0 to 12. | JRC |

## 2.4 Algorithm Outputs

The DSWx-S1 algorithm produces four different output products in cloud optimized GeoTIFF (COG) format. The first output product, WTR, contains all classified layers including water, non-water, high-backscattering water, low-backscattering non-water, land cover mask, hand mask, layover/shadow masks. The second output product, BWTR, contains a binary value including water and non-water and masks. The third output product, CONF, contains the values for the confidence layer computed from the fuzzy logic computation step.

Table 3. Outputs of the DSWx-S1 algorithm.

| Layer Abbreviation | Layer Long Name | Range | Data typeStorage |
|---|---|---|---|
| WTR | Highest processed output layer (classified, land cover, HAND, and layover/shadow masks applied) | 0 - 255 | Int8 |
| BWTR | Binary water map is derived from the WTR layer as a union of water classes | 0 - 255 | Int8 |
| CONF | Confidence layer for WTR | 0 - 255 | Int8 |
| DIAG | Diagnostic layer showing the fuzziness of the water bodies. | 0 - 255 | Int8 |

# 3 ALGORITHM ASSUMPTIONS

- Backscattering of water bodies

The DSWx-S1and DSWx-NI algorithm to estimate open water relies on the assumption that the surface water exhibits the lower backscattering compared to the non-water bodies. The surface water beneath the vegetation canopy induces double bounce scattering and strong returned signals. Thus, this type of water body is beyond our assumption. Furthermore, the water exaction from the ice surface is out of our scope. Since the smooth ice surface has strong specular reflection and similar amplitude characteristics. The algorithm mainly applies to ice-free water bodies.

- Static ancillary data

The DSWx-S1and DSWx-NI algorithms rely on the assumption that the ancillary data used for processing is a static dataset. Any update to the DEM used for processing may impose the reprocessing of the entire DSWx-S1 and DSWx-NI archives.

# 4 ALGORITHM IMPLEMENTATION

The OPERA DSWx-S1 and DSWx-NI product workflow is developed by the OPERA Algorithm Development Team (ADT). It is implemented in Python and it is open-source accessed through the following GitHub repository:

https://github.com/opera-adt/DSWX-SAR

The OPERA RTC-S1 product workflow is developed by the OPERA Algorithm Development Team (ADT). It is implemented in Python and it is open-source accessed through the following GitHub repository:

https://github.com/opera-adt/RTC

# 5 ALGORITHM USAGE CONSTRAINTS

The DSWX-S1and DSWx-NI algorithms does not account for or estimate:

- possible geolocation errors in the input data
- water when or where the vegetation canopy is sufficiently dense to obscure most surface water in the pixel (inundated vegetation)
- liquid water over non-liquid water
- errors due to Radio Frequency Interference (RFI)
- errors due to high-backscattering objects over the water surface.

# 6 PERFORMANCE ASSESSMENT VALIDATION

Algorithm calibration and validation was baselined by the DSWx-S1 product requirements. They specify that for 80% of the validation products considered, overall accuracy must be greater than or equal to 80% for open water given features with a minimum area of 3ha and minimum width of 200m.

## 6.1 Performance Assessment Validation Methods

For DSWx-S1 performance verification, the water extent data generated from the optical images or a different algorithm using Sentinel-1 can be cross-compared.

Table 4. Data for validation

| Input Data | Description | Producer |
|---|---|---|
| DSWE | Generated from Landsat for US regions with 30 m spacing at the time of input image collection | USGS |
| DSWx-HLS | Generated from Harmonized Landsat and Sentinel-2 for global areas with 30 m spacing at the time of input image collection. | JPL |
| GLAD surface water dynamics | individual month water percent generated from Landsat 5,7, and 8 | - |

| | | |
|---|---|---|
| JRC water occurrence | Water occurrence generated from collection of Landsat with values ranging from 0 to 100 | JRC |
| ARIA FPM | Binary layer with flood/no-flood computed from the pair of Sentinel-1 | JPL |

The Dynamic Surface Water Extent (DSWE) product, devised by the USGS, encapsulates surface water extent data derived from Landsat observations. However, it is limited in its geographic scope, covering only the continental US, Alaska, and Hawaii (Jones, 2019). For regions beyond these US territories, DSWx-HLS presents an alternate and comparable source of data. Since both DSWE and DSWx-HLS datasets are generated from individual images at the time of collection, any comparisons drawn must be made with data acquired within a 48-hour window relative to the Sentinel-1 acquisition time. Additionally, to ensure a reliable representation of product performance, data with minimal cloud cover is preferred.

In instances where comparable optical products are unavailable within the specified time frame or acceptable cloud cover thresholds, the GLAD product suite offers an alternative benchmark. GLAD includes a monthly percentage of water extents computed from all available Landsat imagery within a specific month. However, in scenarios where cloud cover persists throughout the month and across all image collection times, certain areas will inevitably return a 'no-data' value. Comparisons involving GLAD should be conducted with caution since it does not delineate detailed water extents for a specific timeframe. Discrepancies between GLAD and DSWx-S1 could emerge due to the inherent characteristics of the data, making this approach primarily suitable for regions with relatively static water extents. For example, the coastline affected by the tidal changes may exhibit the differences in the water extent.

In worst-case scenarios, where none of the aforementioned data is accessible due to pervasive cloud cover, the Joint Research Centre's (JRC) water occurrence map serves as the final resource. This comprehensive map has been compiled from Landsat images acquired over a span of 37 years. Here, the pixels are likely to be filled with the values rather than ‘no-data’. However, caution is required in areas where water extents experience frequent and substantial changes, as the comparison with this static, long-term dataset might prove insufficient.

In the context of a flood event, detecting water underneath cloud cover using optical images presents a significant challenge. Instead of relying solely on optical images for validation, the ARIA Flood Proxy Map (FPM) products can be instrumental. The ARIA FPM products are derived from pairs of Sentinel-1 images captured at pre-flood and co-flood states. Consequently, these FPMs primarily highlight changes in backscatter resulting from floods rather than defining standing water regions. Therefore, regions newly inundated can be accurately assessed using ARIA FPM products. However, it's essential to note that the ARIA FPM is an on-demand product and may not be available for all flood events.

In order to quantify the accuracy of the DSWx-S1, first metric is an ‘overall accuracy’, that is the number of correct classification predictions divided by the total number of classification attempts :

Accuracy = (TP + TN) / (TP + FN + FP + TN) where

TP = True Positives (class in question is shown where present)

TN = True Negatives (absence of class in question is shown where absent)

FP = False Positives (class in questions is incorrectly shown as present)

FN = False Negatives (absence of class is shown where it is actually present)

However, OA is less informative when mapped classes are imbalanced. In such cases, the metrics listed in the table can provide complementary insights into evaluating the performance.

Table 4. Data for validation

| Metric | Definition | Formula | Notes |
|---|---|---|---|
| Accuracy | Of the total number of evaluation points, how many were correctly classified? | (TP + TN) /<br>(TP + FN + FP + TN) | OPERA Requirement. Equal sampling provides conservative accuracy estimate |
| **P**recision | Of the total positive predictions, what number were correct? | TP /<br>(TP + FP) | Not a Requirement<br>Ignores (overwhelming) ‘not water’ correct predictions |
| **R**ecall | Of the total number of evaluation points, how many positive predictions were correct? | TP /<br>(TP + FN) | Not a Requirement |
| Kappa coefficient | Agreement of binary classification | $\frac{2 \times (T_p \times T_n -}{(T_p + F_p) \times (T_n + F_p) + (}$ | Provides conservative accuracy estimate |

## 6.2 Performance Assessment Validation Uncertainties

If a difference in acquisition times between the reference image and DSWx products computed from Sentinel-1 and NISAR exist, discrepancies in water extent observations may arise. In regions where static water bodies predominate, accuracy metrics such as overall accuracy, user/producer accuracy, and kappa statistics tend to be less influenced by the timing of the acquisition. Conversely, in coastal areas where tidal water level changes significantly impact water extents, these differences in acquisition times can lead to more substantial variations in the observed water extents. This discrepancy can happen even when the difference of the acquisition times is less than 1 day.

## 6.3 Performance Assessment Validation Errors

If the power of RFI is too strong, it degrades image quality by making image blurring and showing bright lines. When the Sentinel-1 Image is affected by the RFI, the bright lines can be interpreted as the non-water bodies in the proposed algorithm.

Similarly, the strong backscattering objects are the sources hampering the surface water detection. The strong backscattering typically has the strong sidelobes. If these types of targets are located near the water bodies, their sidelobes are relatively bright over low-backscattering water and may underestimate the water extent. Since boats and ships are strong scatterers, some areas surrounding these objects could be misclassified as the non-water pixels. In the extreme cases, the cluster of the ships having strong backscattering signals may affect the algorithm performance by showing huge false negative errors.

# 7 DATA ACCESS

Data Access Input Data

Data Access Output Data:

- DSWx-S1 products will be made available through the National Aeronautics and Space Administration (NASA) Physical Oceanography Distributed Active Archive Center (PO.DAAC).

Data Access Related URLs

# 8 CONTACTS

Heresh Fattahi
OPERA Algorithm Development Team
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, CA, USA
heresh.fattahi@jpl.nasa.gov

Jungkyo Jung
OPERA Algorithm Development Team
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, CA, USA
jungkyo.jung@jpl.nasa.gov

# 9 REFERENCES

Alsdorf, D.E., Rodríguez, E., Lettenmaier, D.P., 2007. Measuring surface water from space. Rev. Geophys. 45.

Ashman, K.M., Bird, C.M., Zepf, S.E., 1994. Detecting bimodality in astronomical datasets. ArXiv Prepr. Astro-Ph9408030.

Bauer-Marschallinger, B., Cao, S., Navacchi, C., Freeman, V., Reuß, F., Geudtner, D., Rommen, B., Vega, F.C., Snoeij, P., Attema, E., 2021. The normalised Sentinel-1 Global Backscatter Model, mapping Earth's land surface with C-band microwaves. Sci. Data 8, 1–18.

Brakenridge, R., Anderson, E., 2006. MODIS-based flood detection, mapping and measurement: the potential for operational hydrological applications, in: Transboundary Floods: Reducing Risks through Flood Management. Springer, pp. 1–12.

Brisco, B., 2015. Mapping and monitoring surface water and wetlands with synthetic aperture radar. Remote Sens. Wetl. Appl. Adv. 119–136.

Chini, M., Hostache, R., Giustarini, L., Matgen, P., 2017. A hierarchical split-based approach for parametric thresholding of SAR images: Flood inundation as a test case. IEEE Trans. Geosci. Remote Sens. 55, 6975–6988.

Clement, M.A., Kilsby, C.G., Moore, P., 2018. Multi-temporal synthetic aperture radar flood mapping using change detection. J. Flood Risk Manag. 11, 152–168.

DeVries, B., Huang, C., Lang, M.W., Jones, J.W., Huang, W., Creed, I.F., Carroll, M.L., 2017. Automated quantification of surface water inundation in wetlands using optical satellite imagery. Remote Sens. 9, 807.

Ding, X., Li, X., 2011. Monitoring of the water-area variations of Lake Dongting in China with ENVISAT ASAR images. Int. J. Appl. Earth Obs. Geoinformation 13, 894–901.

Du, Z., Li, W., Zhou, D., Tian, L., Ling, F., Wang, H., Gui, Y., Sun, B., 2014. Analysis of Landsat-8 OLI imagery for land surface water mapping. Remote Sens. Lett. 5, 672–681.

Elkhrachy, I., 2018. Assessment and management flash flood in Najran Wady using GIS and remote sensing. J. Indian Soc. Remote Sens. 46, 297–308.

Ellison, A.M., 1987. Effect of seed dimorphism on the density-dependent dynamics of experimental populations of Atriplex triangularis (Chenopodiaceae). Am. J. Bot. 74, 1280–1288.

García-Pintado, J., Mason, D.C., Dance, S.L., Cloke, H.L., Neal, J.C., Freer, J., Bates, P.D., 2015. Satellite-supported flood forecasting in river networks: A real case study. J. Hydrol. 523, 706–724.

Giustarini, L., Hostache, R., Matgen, P., Schumann, G.J.-P., Bates, P.D., Mason, D.C., 2012. A change detection approach to flood mapping in urban areas using TerraSAR-X. IEEE Trans. Geosci. Remote Sens. 51, 2417–2430.

Gong, M., Li, H., Jiang, X., 2016. A multi-objective optimization framework for ill-posed inverse problems. CAAI Trans. Intell. Technol. 1, 225–240.

Hostache, R., Chini, M., Giustarini, L., Neal, J., Kavetski, D., Wood, M., Corato, G., Pelich, R.-M., Matgen, P., 2018. Near-real-time assimilation of SAR-derived flood maps for improving flood forecasts. Water Resour. Res. 54, 5516–5535.

Jones, J.W., 2019. Improved automated detection of subpixel-scale inundation—Revised dynamic surface water extent (DSWE) partial surface water tests. Remote Sens. 11, 374.

Kellndorfer, J., Cartus, O., Lavalle, M., Magnard, C., Milillo, P., Oveisgharan, S., Osmanoglu, B., Rosen, P.A., Wegmüller, U., 2022. Global seasonal Sentinel-1 interferometric coherence and backscatter data set. Sci. Data 9, 1–16.

Khandelwal, A., Karpatne, A., Marlier, M.E., Kim, J., Lettenmaier, D.P., Kumar, V., 2017. An approach for global monitoring of surface water extent variations in reservoirs using MODIS data. Remote Sens. Environ. 202, 113–128.

Kittler, J., Illingworth, J., 1986. Minimum error thresholding. Pattern Recognit. 19, 41–47.

Lesack, L.F., Melack, J.M., 1995. Flooding hydrology and mixture dynamics of lake water derived from multiple sources in an Amazon floodplain lake. Water Resour. Res. 31, 329–345.

Li, Y., Niu, Z., Xu, Z., Yan, X., 2020. Construction of high spatial-temporal water body dataset in China based on Sentinel-1 archives and GEE. Remote Sens. 12, 2413.

Liang, J., Liu, D., 2020. A local thresholding approach to flood water delineation using Sentinel-1 SAR imagery. ISPRS J. Photogramm. Remote Sens. 159, 53–62.

Liao, P.-S., Chen, T.-S., Chung, P.C., 2001. A Fast Algorithm for Multilevel Thresholding. J. Inf. Sci. Eng. 17, 713–727.

Lin, Y.N., Yun, S.-H., Bhardwaj, A., Hill, E.M., 2019. Urban flood detection with Sentinel-1 multi-temporal synthetic aperture radar (SAR) observations in a Bayesian framework: a case study for Hurricane Matthew. Remote Sens. 11, 1778.

Liu, H., Jezek, K.C., 2004. Automated extraction of coastline from satellite imagery by integrating Canny edge detection and locally adaptive thresholding methods. Int. J. Remote Sens. 25, 937–958.

Martinis, S., Kersten, J., Twele, A., 2015. A fully automated TerraSAR-X based flood service. ISPRS J. Photogramm. Remote Sens. 104, 203–212.

Martinis, S., Plank, S., Ćwik, K., 2018. The use of Sentinel-1 time-series data to improve flood monitoring in arid areas. Remote Sens. 10, 583.

Martinis, S., Twele, A., Voigt, S., 2009. Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data. Nat. Hazards Earth Syst. Sci. 9, 303–314.

Matgen, P., Hostache, R., Schumann, G., Pfister, L., Hoffmann, L., Savenije, H.H.G., 2011. Towards an automated SAR-based flood monitoring system: Lessons learned from two case studies. Phys. Chem. Earth Parts ABC 36, 241–252.

Merz, B., Kreibich, H., Schwarze, R., Thieken, A., 2010. Review article" Assessment of economic flood damage". Nat. Hazards Earth Syst. Sci. 10, 1697–1724.

Messner, F., Meyer, V., 2006. Flood damage, vulnerability and risk perception–challenges for flood damage research, in: Flood Risk Management: Hazards, Vulnerability and Mitigation Measures. Springer, pp. 149–167.

Nakmuenwai, P., Yamazaki, F., Liu, W., 2017. Automated extraction of inundated areas from multi-temporal dual-polarization RADARSAT-2 images of the 2011 central Thailand flood. Remote Sens. 9, 78.

Ouled Sghaier, M., Hammami, I., Foucher, S., Lepage, R., 2018. Flood extent mapping from time-series SAR images based on texture analysis and data fusion. Remote Sens. 10, 237.

Pekel, J.-F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. Nature 540, 418–422.

Pfister, R., Janczyk, M., 2013. Confidence intervals for two sample means: Calculation, interpretation, and a few simple rules. Adv. Cogn. Psychol. 9, 74.

Psomiadis, E., 2016. Flash flood area mapping utilising SENTINEL-1 radar data, in: Earth Resources and Environmental Remote Sensing/GIS Applications VII. SPIE, pp. 382–392.

Pulvirenti, L., Pierdicca, N., Chini, M., Guerriero, L., 2013. Monitoring flood evolution in vegetated areas using COSMO-SkyMed data: The Tuscany 2009 case study. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 6, 1807–1816.

Shiroma, G.H., Lavalle, M., Buckley, S.M., 2022. An area-based projection algorithm for SAR radiometric terrain correction and geocoding. IEEE Trans. Geosci. Remote Sens. 60, 1–23.

Small, D., 2011. Flattening gamma: Radiometric terrain correction for SAR imagery. IEEE Trans. Geosci. Remote Sens. 49, 3081–3093.

Tay, C.W., Yun, S.-H., Chin, S.T., Bhardwaj, A., Jung, J., Hill, E.M., 2020. Rapid flood and damage mapping using synthetic aperture radar in response to Typhoon Hagibis, Japan. Sci. Data 7, 1–9.

Twele, A., Cao, W., Plank, S., Martinis, S., 2016. Sentinel-1-based flood mapping: a fully automated processing chain. Int. J. Remote Sens. 37, 2990–3004.

Uddin, K., Matin, M.A., Meyer, F.J., 2019. Operational flood mapping using multi-temporal Sentinel-1 SAR images: A case study from Bangladesh. Remote Sens. 11, 1581.

Wieland, M., Martinis, S., 2020. Large-scale surface water change observed by Sentinel-2 during the 2018 drought in Germany. Int. J. Remote Sens. 41, 4742–4756.

Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., 2021. ESA WorldCover 10 m 2020 v100. Zenodo Geneve Switz.

Zhao, J., Pelich, R., Hostache, R., Matgen, P., Cao, S., Wagner, W., Chini, M., 2021. Deriving exclusion maps from C-band SAR time-series in support of floodwater mapping. Remote Sens. Environ. 265, 112668.

# 10 ACRONYMS

| | |
|---|---|
| ADT | Algorithm Development Team |
| ARIA | Advanced Rapid Imaging and Analysis |
| ASF | Alaska Satellite Facility |
| DAAC | Distributed Active Archive Center |
| DEM | Digital Elevation Model |
| DSWx | Dynamic Surface Water Extent |
| ECEF | Earth-centered Earth-fixed |
| ESA | European Space Agency |
| FPM | Flood Proxy Map |
| HAND | Height Above Nearest Drainage |
| HLS | Harmonized Landsat Sentinel-2 |
| JPL | Jet Propulsion Laboratory |
| JRC | Joint Research Centre |
| L1 | Level-1 product |
| L2 | Level-2 product |
| NASA | National Aeronautics and Space Administration |
| NISAR | NASA-ISRO SAR |
| OPERA | Operational Product for End-user from Remote sensing Analysis |
| RFI | Radio Frequency Interference |
| RTC | Radiometric Terrain Corrected |
| S1 | Sentinel-1 |
| S1-A/B | Sentinel-1 A/B |
| SAR | Synthetic Aperture Radar |
| SLC | Single-Look Complex |
| UAVSAR | Uninhabited Aerial Vehicle Synthetic Aperture Radar |
| WGS84 | World Geodetic System 1984 |

# 11 LIST OF SYMBOLS

| | |
|---|---|
| $\mu$ | Mean |
| $\sigma$ | Standard deviation |

# 12 ACKNOWLEDGEMENTS

This research was conducted at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

The original Copernicus Sentinel data used for this paper have been provided by the European Space Agency.