



**fire**  
cci

---


## ESA Climate Change Initiative – Fire\_cci

### D2.1.1 Algorithm Theoretical Basis Document: SAR-Optical

---

<b>Project Name</b>	ECV Fire Disturbance: Fire_cci
<b>Contract N°</b>	4000126706/19/I-NB
<b>Issue Date</b>	27/04/2020
<b>Version</b>	1.1
<b>Author</b>	Miguel Angel Belenguer-Plomer and Mihai A. Tanase
<b>Document Ref.</b>	Fire_cci_D2.1.1_ATBD_SAR-O_1.1
<b>Document type</b>	Public

*To be cited as: Belenguer-Plomer M.A. and Tanase M.A. (2020) ESA CCI ECV Fire Disturbance: D2.1.1 SAR-Optical Algorithm Theoretical Basis Document: version 1.1. Available at: <https://climate.esa.int/en/projects/fire/key-documents/>*

	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1		
		Issue	1.1	Date	27/04/2020
				Page	2


## Project Partners

Prime Contractor/ Scientific Lead & Project Management	UAH – University of Alcalá (Spain)
Earth Observation Team	UAH – University of Alcalá (Spain)
	UPM – Universidad Politécnica de Madrid (Spain)
	CNR-IREA - National Research Council of Italy – Institute for Electromagnetic Sensing of the Environment (Italy)
System Engineering	BC – Brockmann Consult (Germany)
Climate Modelling Group	MPIM – Max Planck Institute for Meteorology (Germany)
	CNRS - National Centre for Scientific Research (France)



## Distribution

Affiliation	Name	Address	Copies
ESA	Simon Pinnock (ESA)	simon.pinnock@esa.int	electronic copy
	Clement Albergel (ESA)	clement.albergel@esa.int	
Project Team	Emilio Chuvieco (UAH)	emilio.chuvieco@uah.es	electronic copy
	M. Lucrecia Pettinari (UAH)	mlucares.pettinari@uah.es	
	Joshua Lizundia (UAH)	joshua.lizundia@uah.es	
	Gonzalo Otón (UAH)	gonzalo.oton@uah.es	
	Mihai Tanase (UAH)	mihai.tanase@uah.es	
	Miguel Ángel Belenguer (UAH)	miguel.belenguer@uah.es	
	Consuelo Gonzalo (UPM)	consuelo.gonzalo@upm.es	
	Dionisio Rodríguez Esparragón (UPM)	dionisio.rodriquez@ulpgc.es	
	Ángel García Pedrero (UPM)	angelmario.garcia@upm.es	
	Daniela Stroppiana (CNR)	stroppiana.d@irea.cnr.it	
	Mirco Boschetti (CNR)	boschetti.m@irea.cnr.it	
	Thomas Storm (BC)	thomas.storm@brockmann-consult.de	
	Martin Böttcher (BC)	martin.boettcher@brockmann-cons...	
Grit Kirches (BC)	grit.kirches@brockmann-consult.de		
Angelika Heil (MPIM)	angelika.heil@mpimet.mpg.de		
Idir Bouarar (MPIM)	idir.bouarar@mpimet.mpg.de		
Florent Mouillot (CNRS)	florent.mouillot@cefe.cnrs.fr		
Philippe Ciais (CNRS)	<a href="mailto:philippe.ciais@lsce.ipsl.fr">philippe.ciais@lsce.ipsl.fr</a>		

	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1		
		Issue	1.1	Date	27/04/2020
				Page	3

## Summary

This document describes the algorithms used for generating the small-fire dataset for the three demonstrator areas selected in Sub-Sahara South Africa within the Fire\_cci project. It includes the description of the Sentinel-1 C-band data pre-processing, as well as the ancillary data used to derive burned area over the tropical Africa.

	Affiliation/Function	Name	Date
<b>Prepared</b>	UAH	Mihai A. Tanase	27/04/2020
	UAH	Miguel Angel Belenguer	
<b>Reviewed</b>	UAH – Project Manager	Lucrecia Pettinari	27/04/2020
<b>Authorized</b>	UAH - Science Leader	Emilio Chuvieco	27/04/2020
<b>Accepted</b>	ESA - Technical Officer	Clément Albergel	05/05/2020

This document is not signed. It is provided as an electronic copy.

## Document Status Sheet

Issue	Date	Details
1.0	30/03/2020	First release of the document.
1.1	27/04/2020	Addressing comments of Fire_cci+_D2.2.1_ATBD_SAR-O_v1.0_RID.doc


## Document Change Record

Issue	Date	Request	Location	Details
1.1	27/04/2020	ESA	Summary Executive Summary Section 2 Sections 2.4, 3.2, 3.3, 3.3.1, 3.3.2, 3.3.3, 4.2, 4.4, 4.6.2, 4.6.3, 4.7.1, 4.8 Section 4.5 Section 5	Text updated Text updated, and partly moved to Section 2 Information reorganized and updated Small changes in the text  Text expanded References updated



## **Table of Contents**

<b>1. Executive Summary</b> .....	<b>6</b>
<b>2. Introduction</b> .....	<b>6</b>
2.1. Background.....	6
2.2. Purpose of the document .....	7
2.3. Applicable Documents .....	7
2.4. The SAR-Optical algorithm: general considerations.....	7
<b>3. Data</b> .....	<b>8</b>
3.1. Sentinel-1 system and data products .....	8
3.2. Sentinel-2 system and data products .....	9
3.3. Ancillary datasets .....	10
3.3.1. Landsat-8 data products .....	10
3.3.2. Active fire products.....	11
3.3.3. Other datasets.....	11
<b>4. Methods</b> .....	<b>12</b>
4.1. Site selection for algorithm development and preliminary validation .....	12
4.2. Sentinel-1 pre-processing .....	12
4.3. Sentinel-2 pre-processing .....	13
4.4. SAR-optical data integration .....	14
4.5. Reference burned perimeters and validation .....	15
4.6. Burned area mapping.....	16
4.6.1. Convolutional Neural Networks (CNN) - background.....	16
4.6.2. Selection of training data .....	17
4.6.3. Assessment of optimum CNN configuration for BA mapping.....	17
4.7. Preliminary results and significance for BA mapping algorithm development ...	18
4.7.1. Optimum CNN configuration .....	19
4.7.2. SAR-optical mapping strategy .....	23
4.8. Preliminary algorithm validation.....	24
<b>5. References</b> .....	<b>27</b>
<b>Annex 1: Acronyms and abbreviations</b> .....	<b>31</b>


	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1		
		Issue	1.1	Date	27/04/2020
				Page	5

## **List of Tables**

Table 1: Main characteristics of the Sentinel-1 satellite sensor. ....	9
Table 2: Band characteristics of the Sentinel-2 MSI.....	9
Table 3: Landsat 8 OLI and TIRS bands description .....	10
Table 4: Landsat 8 and Sentinel-1/2 data used for algorithm development and preliminary validation .....	12
Table 5: Confusion matrix (example).....	16
Table 6: The eight configurations assessed for each input dataset (S – simple, C – complex).....	18
Table 7: Optimum CNN configuration and the average Dice Coefficient of the calibration tiles (DC $\pm$ standard deviation) achieved by land cover class and input data .....	22
Table 8: Error metrics for Sentinel-1 burned area detections for each MGRS tile analysed .....	25

## **List of Figures**

Figure 1: Comparison of Landsat and Sentinel 2 spectral bands (source: NASA) .....	11
Figure 2: Areas selected for algorithm development and preliminary validation. Terrestrial eco-regions are also shown. ....	12
Figure 3: Flowchart for SAR data processing with Orfeo Toolbox .....	13
Figure 4: Central wavelength (nm) by bands number of the MSI onboard Sentinel-2 A and B, and Graphical representation of temporal composite formation.....	14
Figure 5: Configuration of 1D (left) and 2D (right) CNNs (Belenguer-Plomer et al., in preparation).....	18
Figure 6: Dice coefficient (DC), commission and omission errors (CE and OE) by calibration tiles when training different CNN configuration and input data. ....	20
Figure 7: Mean and standard error of Dice coefficient (DC), commission and omission errors (CE and OE) by land cover classes when training different CNN configuration and input data. O: other. F: forest. S: shrubland. G: grassland. C: cropland.....	21
Figure 8: Variation of mapping accuracy as a function of changes in softmax probability by land cover class and input data sets .....	22
Figure 9: Mean and standard error of Dice coefficient (DC), commission and omission errors (CE and OE) by land cover classes when combining Sentinel-1 and -2 applying four different approaches, being a - a SAR and optical data integration before the CNN based classification; b – the Sentinel-2 BA based maps pixels affected by no data pixels were filled using BA from Sentinel-1; and c - a SAR and optical integration of all burned pixels. ....	24
Figure 10: BA mapping from active (S1), passive (S2) and active-passive (S1+S2) sensors .....	26
Figure 11: BA mapping accuracy for the validating tiles as a function of the sensor used .....	26

	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1			
		Issue	1.1	Date	27/04/2020	
					Page	6

## 1. Executive Summary

This Algorithm Theoretical Basis Document (ATBD), corresponding to the SAR-O algorithm, describes the algorithm, methods, and approaches that lead to the generation of a small fire dataset for three 400x400 km areas of interest (AOI) in Africa. Test sites for algorithm development and calibration have been selected based on representative vegetation types (i.e., grasslands, crops and forests) that burn on a regular basis. The theoretical basis described here identifies the data sets used to classify burned area and the methods used to derive the cartographic products. Burned area mapping products derived from Sentinel-1/2 imagery were created for the three AOIs for the year 2019. The resulting maps will be independently validated using BA perimeters derived from multispectral optical datasets (e.g., Sentinel-2, Landsat-8, PLANET).

## 2. Introduction


### 2.1. Background

The European Space Agency (ESA) Climate Change Initiative (CCI) initiative stresses the importance of providing a higher scientific visibility to data acquired by ESA sensors, especially in the context of the Intergovernmental Panel on Climate Change (IPCC) reports. This implies producing consistent time series of accurate Essential Climate Variables products, which can be used by the climate, atmospheric and ecosystem scientists. The importance of keeping long-term observations and the international links with other agencies currently generating ECV data is also stressed.

Fires influence greenhouse gases budget, with a causal relationship between biomass burning and inter-annual variability of related emissions being observed (Simmonds et al. 2005). Worldwide, about 350 million hectares are affected by large fires (i.e., mapped from 500 m spatial resolution data) annually (Giglio et al. 2013), exerting a major influence on carbon release from terrestrial ecosystems (Andreae and Merlet 2001; Simmonds et al. 2005). Fires are also a major factor in land cover changes, and hence affect fluxes of energy and water to the atmosphere. In this context, spatial and temporal monitoring of burned areas can be inferred using remote sensing, a cost effective, objective, and time-saving method to monitor and quantify location, extent and intensity of fire events (Chuvieco 1999; Laneve et al. 2006; Stroppiana et al. 2003). The Fire Disturbance Essential Climate Variable (ECV) provides baseline products to allow such monitoring activities at global scales.

This ECV identifies burned area (BA) as the primary fire variable. Accordingly, the Fire\_cci project focuses on developing and validating algorithms to meet Global Climate Observing System (GCOS) ECV requirements for (consistent, stable, error-characterised) global satellite data products from multi-sensor data archives. Burned Area is defined as any vegetated area that has been completely or partially consumed by a fire, regardless of whether that fire was of human or natural origin, or whether that fire affected wildland areas or human managed territories (agricultural or pastures).

BA can be combined with information on burn efficiency and available fuel load to estimate emissions of trace gases and aerosols. Measurements of BA may be used as direct input (driver) to climate and carbon cycle models or, when long time series of data are available, to parameterize climate-driven models for BA. Even though most destructive fires are large, small fires (<100 ha) may have a relevant contribution on

	<b>fire</b> cci	Fire_cci		Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1		
		SAR-Optical Algorithm		Issue	1.1	Date	27/04/2020
		Theoretical Basis Document				Page	7

atmospheric emissions and carbon budgets (van der Werf et al. 2017). However, precise assessments of that contribution are not yet available. Model based studies estimated that small fires may account for 35% of global burned area (Randerson et al. 2012).

Global BA products are based on coarse resolution sensors (from 250 to 1000m). Since coarse resolution burned area detection algorithms require that a substantial fraction of an individual pixel's area undergo burning for successful attribution (to avoid commission errors from other forms of land cover change), detection of small fires becomes difficult. Therefore, the likelihood of detecting small burns (i.e. < 50ha) is low, and the products are frequently affected by omission errors (Giglio et al. 2009; Padilla et al. 2015), particularly coming from small fires (Kloster et al. 2012). To improve the characterization of small fires, the Fire\_cci project generated a small fire database (SFD) based on medium resolution sensors (10 to 100m). Considering the massive processing effort when generating products at global level, the SFD was focused on the African continent (Roteta et al. 2019), the most burned worldwide (Chuvienco et al. 2016; Giglio et al. 2013), with additional areas being *a posteriori* selected over tropical South-east Asia (Lohberger et al. 2017) and South America (Belenguer-Plomer et al. 2019b).

The aim of WP2100 (Fire\_cci+ Phase 1) is to better understand the impact of different sensors for the detection of burned areas. The combined SAR-optical (SAR-O) BA algorithm was developed to seamlessly integrate information from different sensor types at pixel level. The algorithm integrates pixel-based approaches with object recognition and contextual information. The algorithm detects changes within optical indices and/or SAR backscatter coefficient between consecutive periods and relates them to burned areas by training a convolution neural network (CNN) (LeCun et al. 2015; Zhu et al. 2017). The iterative multi-temporal analysis takes advantage of ancillary information on land cover and thermal anomalies (hotspots) to label the fire affected areas used for CNN training.

## 2.2. Purpose of the document

This document provides the description of the SAR-O algorithm together with preliminary validation results over five Military Grid Reference System (MGRS) tiles distributed worldwide. The document draws from the ATBDs already provided within Fire\_cci Phase 2.


## 2.3. Applicable Documents

[RD-1]	Bastarrika A., Roteta E. (2018) ESA CCI ECV Fire Disturbance: D2.1.2 Algorithm Theoretical Basis Document-SFD, version 1.0. Available at: <a href="https://www.esa-fire-cci.org/documents">https://www.esa-fire-cci.org/documents</a>
[RD-2]	M.A. Tanase, M.A. Belenguer-Plomer (2018) ESA CCI ECV Fire Disturbance: O3.D1 Algorithm Theoretical Basis Document – S1 South America, v2.0. Available at: <a href="https://www.esa-fire-cci.org/documents">https://www.esa-fire-cci.org/documents</a>

## 2.4. The SAR-Optical algorithm: general considerations

The growing availability of medium resolution (10-50 m) optical sensors, such as Sentinel-2 Multispectral Instrument (MSI) and Landsat Operational Land Imager (OLI), has opened new opportunities to characterize the impact of small fires, although their use is limited by the persistent cloud cover in tropical regions. SAR data provide the means



	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1			
		Issue	1.1	Date	27/04/2020	
					Page	8

to overcome the drawbacks of the optical sensors in areas with persistent cloud cover. Over the past decade, radar sensors have become of interest for tracking fire disturbance due to their independence of cloud cover and solar illumination (Bourgeau-Chavez et al. 2002; Tanase et al. 2010). The standard derivative from radar data is the backscatter coefficient which provides an indication of the amount of returned energy that is scattered in a backward direction towards the sensor. *A priori*, the removal of vegetative material reduces the number of scattering elements and thus the total backscatter, as elements of a size similar with the wavelength constitute important sources for microwave scattering (Rignot and Zyl 1993).

The combined use of optical and SAR sensors may reduce omission errors over areas with frequent cloud cover at the cost of making global processing computationally demanding. Indeed, despite the increasing processing power of cloud computing systems generation of continental BA products from Sentinel-2 sensors requires significant effort. Therefore, a combined SAR-Optical algorithm needs to efficiently integrate and use the two data types. Such efficiency may be achieved by dynamic inclusion of SAR datasets into the processing chains depending on pixel-wise availability of optical information over the selected detection period. Further, computational efficiency may be achieved by using state of the art algorithms such as deep learning (DL). DL has been applied in many remote sensing-based studies over the last years and constitutes an efficient way of integrating multiple data sources. Among the deep learning methods, convolutional neural networks (CNN) are widely used in the remote sensing field and has been selected as the basis for the development of the optical-radar algorithm (Ban et al. 2020; Pinto et al. 2020).


As with all change detection methods, it is essential to link the change in optical reflectance and/or backscatter coefficient to fire activity. Contextual information data may help reduce uncertainties by taking advantage of information from the temporal (e.g., higher fire probabilities for certain months) or spatial (e.g., burned area size, shape and distance to neighbours) domains. Furthermore, high resolution active fire products (e.g., Visible Infrared Imaging Radiometer Suite -VIIRS sensor, aboard the Suomi National Polar-orbiting Partnership – NASA (NPP) satellite) may be linked to changed patches based on location and acquisition time and thus provide additional means when identifying areas affected by relatively small fires.

### 3. Data

#### 3.1. Sentinel-1 system and data products

Sentinel-1 is a two-satellite constellation (A - since April 2014, B - since April 2016) with the prime objectives of Land and Ocean monitoring (Table 1). The satellites carry a C-band SAR sensor, which offers medium and high-resolution imaging in all weather conditions making it useful for land monitoring. The radar operates in two main modes, with the Interferometric Wide (IW) swath (250 km width) being the default operation mode over land. The IW mode images three sub-swathes using the Terrain Observation with Progressive Scans SAR (TOPSAR) to provide high quality, homogeneous images. The advantages of Sentinel-1 sensor over other C-band SAR missions, besides the free data access policy, are three-fold: i) high temporal frequency (6 days exact repeat cycle with two satellites), ii) high spatial resolution (5 m in azimuth and 20 m in range) and, iii) dual-polarization (VV and VH).



	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1			
		Issue	1.1	Date	27/04/2020	
					Page	9

Sentinel-1 products are released in two Level 1 formats, Ground Range Detected (GRD) and Single Look Complex (SLC). GRD products are projected, intensity images, radiometrically and terrain corrected. SLC data are designed for interferometric applications, containing both phase and intensity information.

The SAR-O algorithm uses the systematically distributed Level-1 Ground Range Detected, focused data that has been detected, multi-looked and projected to ground range using an Earth ellipsoid model. Pixel values represent detected magnitude (i.e., no phase information). The ellipsoid projection of the GRD products is corrected using the terrain height (as specified in the product general annotation) which varies in azimuth but is constant in range. The products are delivered as GeoTiffs together with ancillary files containing information about orbit, noise and calibration.

**Table 1: Main characteristics of the Sentinel-1 satellite sensor.**

Mission and instrument	Characteristics
Sentinel-1 (A/B)	<b>Orbit:</b> near-polar, sun-synchronous, 180° orbit phasing between A and B <b>Repeat-pass:</b> 12 days per satellite. Combined A/B satellite passes: 6 days. Combined A/B satellites Ascending/Descending passes: 3 days. <b>Instrument:</b> C-band (5.3 GHz) synthetic aperture radar <b>Acquisition modes:</b> Strip map (SM) – on demand Extra Wide Swath (EW) – basic operation mode over sea/ocean Interferometric Wide Swath (IW) – basic operation mode over land <b>Spatial resolution (IW):</b> range: 5 m / azimuth: 20 m <b>Swath width (IW):</b> 250 km

### 3.2. Sentinel-2 system and data products

The data used for the SAR-O algorithm are Sentinel-2 MSI Level-1C product, which contains Top-of-atmosphere reflectance projected in UTM WGS84 system. Bottom-of-atmosphere reflectance and a Scene Classification (SCL) were generated using the ESA toolbox. The Sentinel-2 mission allows a high revisit frequency (5 days at the equator with the Sentinel 2A and 2B satellites). The MSI sensor provides a unique combination of high spatial resolution (which varies from 10 m to 60 m), wide field of view (290 km) and spectral coverage, with 13 spectral bands spanning from the visible and the near infrared to the short-wave infrared (Table 2).

**Table 2: Band characteristics of the Sentinel-2 MSI**

MSI Band Number	Centre Wavelength (nm)	Bandwidth (nm)	Resolution (m)
1	443	20	60
2	490	65	10
3	560	35	10
4	665	30	10
5	705	15	20
6	740	15	20
7	783	20	20
8	842	115	10
8a	865	20	20

MSI Band Number	Centre Wavelength (nm)	Bandwidth (nm)	Resolution (m)
9	945	20	60
10	1380	30	60
11	1610	90	20
12	2190	180	20

### 3.3. Ancillary datasets

The main ancillary datasets used for BA algorithm development and calibration are Landsat 8 imagery (used to derive reference burned perimeters) and the thermal anomalies (hotspots) detected by the Visible Infrared Imaging Radiometer Suite (VIIRS) and Moderate Resolution Imaging Spectroradiometer (MODIS) sensors. In addition, the ESA CCI Land cover v2.0.7 dataset (ESA 2017) at 300 m spatial resolution is used to provide information on land cover types and non-burnable areas (e.g., water bodies).

#### 3.3.1. Landsat-8 data products

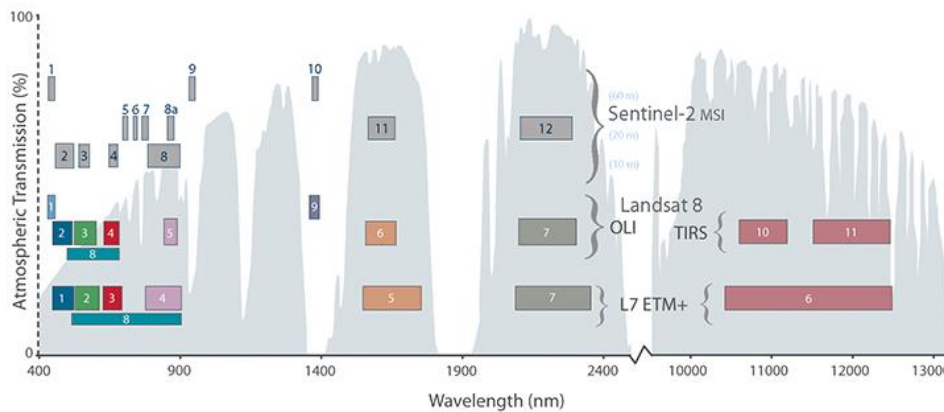
For back-comparability, the SAR-O algorithm development was processed over 10 MGRS tiles used during cal/val activities of the SAR-based Reed-Xiaoli detector (RXD) in Fire\_cci Phase 2 (Belenguer-Plomer et al. 2019b). Landsat-8 images were used to generate reference burned areas perimeters due to the limited availability of Sentinel-2 imagery at the time.

Landsat-8 was launched on February 2013 as a combined effort of NASA and the United States Geological Survey (USGS). Landsat-8 ensures the continued acquisition and availability of Landsat data. It carries two observation sensors, the OLI and the Thermal InfraRed Sensor (TIRS) collecting data in nine shortwave bands and two longwave thermal bands (Table 3 and Figure 1). Landsat-8 Pre-Collection<sup>1</sup> products were used when generating the reference burned area maps for algorithm development and calibration sites.

**Table 3: Landsat 8 OLI and TIRS bands description**

Bands	Wavelength (µm)	Resolution (m)
Band 1 - Coastal aerosol	0.43 - 0.45	30
Band 2 - Blue	0.45 - 0.51	30
Band 3 - Green	0.53 - 0.59	30
Band 4 - Red	0.64 - 0.67	30
Band 5 - Near Infrared (NIR)	0.85 - 0.88	30
Band 6 - SWIR 1	1.57 - 1.65	30
Band 7 - SWIR 2	2.11 - 2.29	30
Band 8 - Panchromatic	0.50 - 0.68	15
Band 9 - Cirrus	1.36 - 1.38	30
Band 10 - Thermal Infrared (TIRS) 1	10.60 - 11.19	100 * (30)
Band 11 - Thermal Infrared (TIRS) 2	11.50 - 12.51	100 * (30)

<sup>1</sup> In 2016, the USGS began making changes to manage the Landsat archive as a tiered Collection of Landsat data. Since October 2017 Landsat Pre-Collection data are no longer available.



**Figure 1: Comparison of Landsat and Sentinel 2 spectral bands (source: NASA)**

### 3.3.2. Active fire products

The SAR-O BA detection algorithm uses active fire products to attribute changes in surface reflectance and backscatter to fire affected areas. The hotspots detected by the VIIRS sensor, onboard the Suomi-NPP satellite (375 m product) seem the most adequate over the areas selected for algorithm development due to its better spatial resolution and a greater response over fires of relatively small areas. Although the MODIS-acquired thermal anomalies were of significant less quality over the selected areas they are used to complement the VIIRS ones. Therefore, vector files of VIIRS and MODIS hotspots downloadable from the Archive Download Tool (<https://firms.modaps.eosdis.nasa.gov/download/>) are used during various stages of the algorithm. One should notice that VIIRS hotspots are affected by errors which are currently being solved by the Fire Information for Resource Management System (FIRMS) team.

### 3.3.3. Other datasets

The ESA CCI Land cover dataset (300 m spatial resolution) was used to provide information on non-burnable areas (e.g., water bodies) and drive the BA detection and mapping algorithm as a function of land cover type. The land cover CCI product (LC\_cci) is an annual series of global land cover maps at 300 m spatial resolution, covering the 1992-2015 period. The maps were produced using a multi-year and multi-sensor strategy to make use of all suitable data and maximize product consistency. The most recent ESA land cover map (i.e., 2015) was used. To account for the different spatial resolution, this layer was resampled to 40 m, i.e. the spatial resolution used for burned area detection and mapping.

The Shuttle Radar Topography Mission DEM (SRTM DEM) provides the reference for SAR data geocoding. The NASA SRTM provides digital elevation information over 80% of the globe. The data is distributed by USGS and is available for download from the National Map Seamless Data Distribution System or the USGS ftp site. The SRTM DEM was derived from single pass SAR interferometric data acquired in 2000. For algorithm development and satellite data processing the enhanced DEM is used. The enhanced DEM was released worldwide in 2016 at 1 arc-seconds resolution (30 m at equator) in 1<sup>0</sup> tiles.

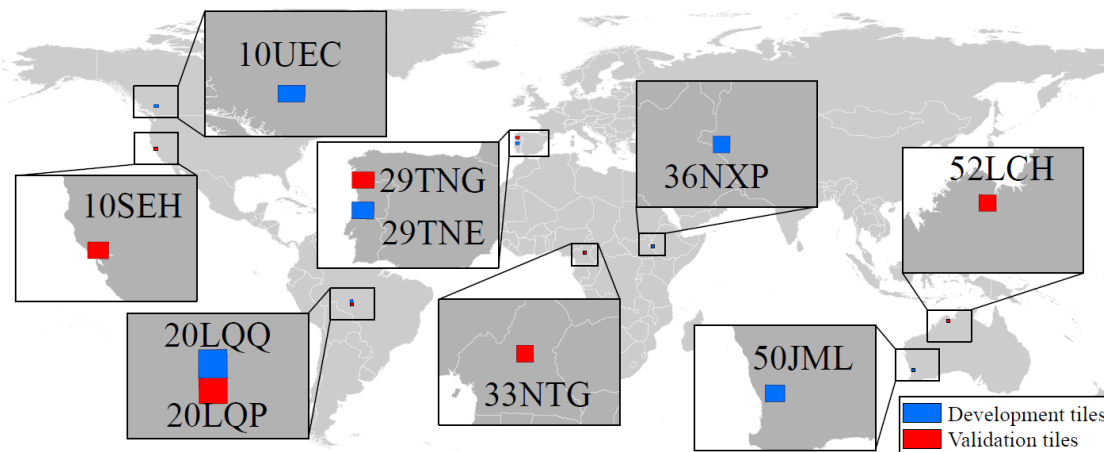
## 4. Methods

### 4.1. Site selection for algorithm development and preliminary validation

Ten areas were selected for algorithm development and preliminary validation<sup>2</sup> (Table 4 and Figure 2) considering: i) the fire activity (from active fire products, i.e., hotspots), ii) the main types of land cover iii) the availability of cloud-free Landsat 8 data. Once the areas for algorithm development and preliminary validation were chosen, burned area perimeters were generated for the 2015-2016 fire seasons, Sentinel-1/2 data were processed, and the BA algorithm was developed, calibrated, and a preliminary validation was carried out.

**Table 4: Landsat 8 and Sentinel-1/2 data used for algorithm development and preliminary validation**

MGRS	Reference period	MGRS	Reference period
10SEH	04/10/17–05/11/17	29TNG	05/10/17–06/11/17
10UEC	05/07/17–22/08/17	33NTG	28/11/15–16/02/16
20LQP	20/07/16–22/09/16	36NXP	30/12/16–15/01/17
20LQQ	04/07/16–22/09/16	50JML	07/03/17–10/05/17
29TNE	05/10/17–06/11/17	52LCH	05/04/17–21/04/17




**Figure 2: Areas selected for algorithm development and preliminary validation. Terrestrial eco-regions are also shown.**

### 4.2. Sentinel-1 pre-processing

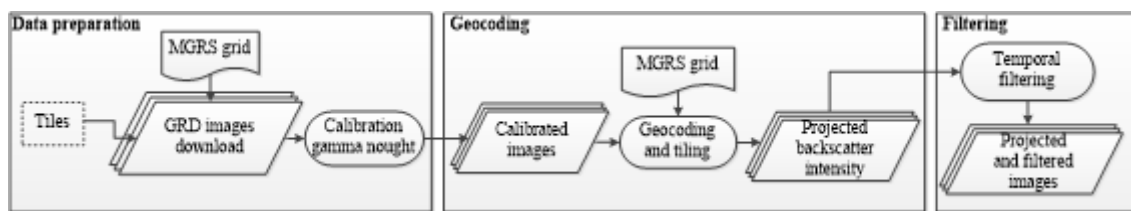
A Sentinel-1 pre-processing chain based on the Orfeo ToolBox (OTB)<sup>3</sup> was selected as it provides for a platform-independent, open source pre-processing solution. The OTB chain was developed by the Centre for the Study of the Biosphere from Space (CESBIO)

<sup>2</sup> The validation of the BA products over the three AOIs will be carried out independently.

<sup>3</sup> OTB is developed by the National Centre for Space Studies (CNES), France

	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1		
		Issue	1.1	Date	27/04/2020
				Page	13

as an operational tool for Sentinel-1 GRD data tiling and processing per the 100 km MGRS used by the Sentinel-2 processing system. The chain is highly scalable (multithreading/multiprocessor) and autonomous once few parameters are set. The chain also deals with data download from the PEPS (Plateforme d'Exploitation des Produits Sentinel) repository that mirrors ESA's Scihub. Alternatively, ESA's Sentinels hub or Copernicus Data and Information Access Services (DIAS) platforms may be used for data download. S1-OTB processing may be grouped in several steps: pre-processing, georeference, and multi-temporal filtering (Figure 3).



**Figure 3: Flowchart for SAR data processing with Orfeo Toolbox**

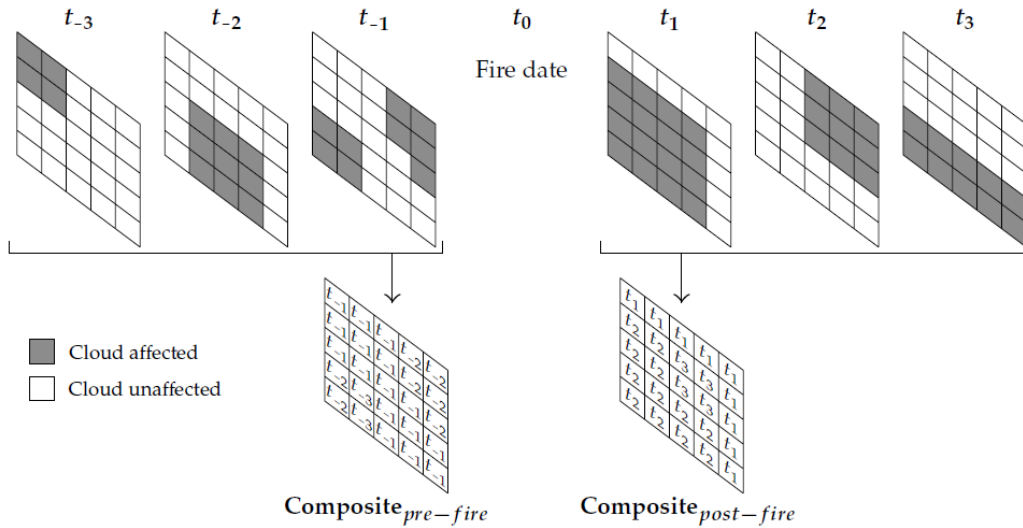
The pre-processing steps include data download (ascending and descending passes) for the specified MGRS tiles, calibration to gamma nought, and multi-look to the desired spatial resolution. The geocoding steps include orthorectification to the desired spatial resolution, subset of Sentinel-1 data to the current processing tile as well as slice assembly for data acquired from the same orbital path but provided within different slices. The last step of the OTB chain is multi-temporal filtering of the products according to satellite pass. Notice that OTB pre-processing chain setup provides images at 20 m spatial resolution. To further reduce speckle and, more importantly, the BA algorithm processing time, the temporally filtered Sentinel-1 images are resampled (i.e. aggregated) to 40 m spatial resolution.

### 4.3. Sentinel-2 pre-processing

The ESA's atmospheric correction algorithm sen2cor (v.2.4.0) was used to derive bottom of atmosphere (BOA) Sentinel-2 images and correct for topographic effects on surface reflectance. The bi-cubic interpolation was subsequently used to resample the 20m Sentinel-2 images to the pre-processed Sentinel-1 output resolution of 40m.

To reduce the number of cloud-affected pixels, temporal composites were generated using both Sentinel-2 A and B images. The sen2cor-based SCL was considered when generating such composites for a given reference date (i.e., pre- or post-fire). Pixels affected by clouds or shadows were gap filled using data from Sentinel-2 imagery acquired at the closest date. Gap filling considered a period of up to 30 days from the reference date. Depending on the reference date, pre-fire or post-fire, the temporal pixel filling was carried using images acquired before or after the reference, respectively (Figure 4).

Satellite	B02	B03	B04	B05	B06	B07	B8a	B11	B12
S2A	492.4	559.8	664.6	704.1	740.5	782.8	864.7	1613.7	2202.4
S2B	492.1	559.0	665.0	703.8	739.1	779.7	864.0	1610.4	2185.7



**Figure 4: Central wavelength (nm) by bands number of the MSI onboard Sentinel-2 A and B, and Graphical representation of temporal composite formation**

Along with surface reflectance for each of the two temporal composites (pre- and post-fire), the following indices were computed and used for BA detection and mapping: NBR, NDVI, NDWI and MIRBI.

$$\text{NBR} = (\text{B8a} - \text{B12}) / (\text{B8a} + \text{B12}) \quad (1)$$

$$\text{NDVI} = (\text{B8a} - \text{B05}) / (\text{B8a} + \text{B05}) \quad (2)$$

$$\text{NDWI} = (\text{B8a} - \text{B11}) / (\text{B8a} + \text{B11}) \quad (3)$$


$$\text{MIRBI} = 10 \times \text{B12} - 9.8 \times \text{B11} + 2 \quad (4)$$

where: B- stands for the band number of Sentinel 2 sensors

#### 4.4. SAR-optical data integration

The burned area detection and mapping algorithm uses time series of Sentinel-1/2 imagery collected with a reasonably short time gap between them. The algorithm is sensitive to the timing of images acquired at  $t-1$  and  $t+1$  which need to be spaced reasonably ( $< 30$  days) as recovering vegetation may obscure the burn signal. Detection and classification are carried out on consecutive acquisition dates for which Sentinel-1 and 2 imagery need to be combined. The SAR-optical data integration is carried out by stacking consecutive pre-fire ( $t-1$ ) and post fire ( $t+1$ ) dates of radar-derived metrics with temporal composites of optical-derived Sentinel-2 metrics (spectral bands). As Sentinel-1 and Sentinel-2 acquisition dates may not coincide, the Sentinel-1 dates are always selected as the detection period due to their complete spatial coverage (e.g., no missing pixels due to cloud cover). Sentinel-2 images are then matched to the Sentinel-1 dates for the current detection period as follows: i) for the pre-fire date, the closest Sentinel-2 image



	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1		
		Issue	1.1	Date	27/04/2020
				Page	15

acquired before (if no coincident image is available) is selected as t-1 date and, ii) for the post-fire date, the closest image acquired after (if no coincident image is available coincident) is selected as t+1 date. Once the Sentinel-2 interval is ‘matched’ with the Sentinel-1 detection interval, the temporal composite process (see Section 4.3) kicks in and cloud related gaps are filled in.

The input pairs of Sentinel-1 metrics and Sentinel-2 image composites for the current detection period are subsequently fed into the classification algorithm and optical indices are derived (see Section 4.3).

#### 4.5. Reference burned perimeters and validation

The reference burned perimeters extraction for validation purposes was based on a well-established framework (Padilla et al. 2014; Padilla et al. 2015; Padilla et al. 2017). The reference data were obtained from Landsat-8 images using a random forest (RF) classifier and training polygons selected by an independent operator. The validation perimeters were generated from multi-temporal pairs of images with a maximum separation of 32 days. The temporal separation of the pairs was short to ensure that fire scars were clearly visible in the post-fire image. Before running the classification, clouds were removed using the pixel quality band of the Landsat product and each pair of images was clipped to the extent of its corresponding MGRS tile. Training areas were selected using a false colour composite (RGB: SWIR, NIR, R) that allowed for a clear discrimination of burned areas. Three training classes were considered: burned, unburned and no data.

The variables selected as input for the RF classifier were: (i) Landsat-8 bands 5 and 7; (ii) the Normalized Burn Ratio (NBR); and (iii) the temporal difference between the pre- and post-fire NBR values (dNBR). The NBR (5) is defined as the normalized difference between the reflectance of NIR and SWIR wavelengths (Key and Benson 2006; López-García and Caselles 1991).

$$NBR = (NIR - SWIR)/(NIR + SWIR) \quad (5)$$

where near infra-red (NIR) has a wavelength of 0.85–0.88  $\mu\text{m}$  and shortwave infra-red (SWIR) has a wavelength of 2.11–2.29  $\mu\text{m}$ .

After the RF classification, fire perimeters were visually revised to correct possible errors. New training fields were iteratively added, and the RF was re-run until the classification results were deemed accurate. Reference BA perimeters were resized using a nearest-neighbour interpolation to the selected pixel spacing of the Sentinel-1 product (40 m). Temporal gaps between the Landsat-8 reference period and the Sentinel-1 detection period were filled in by photointerpretation of Sentinel-2 images.

Confusion matrices were used to validate the burned area maps (Table 5). The Dice coefficient (6), omission errors (7), and commission errors (8), were used to assess the quality of the mapping products obtained within different CNN configurations. For global climate modelling needs, BA products should have commission errors (CE) in the range of 4% (ideal) to 17 % (maximum), with omission errors (OE) above 19% deemed less useful for the modelling efforts (Mouillot et al. 2014).



	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1		
		Issue	1.1	Date	27/04/2020
				Page	16

**Table 5: Confusion matrix (example)**

Detection	Reference data		Row total
	Burned	Unburned	
Burned	$P_{11}$	$P_{12}$	$P_{1+}$
Unburned	$P_{21}$	$P_{22}$	$P_{2+}$
Col. total	$P_{+1}$	$P_{+2}$	$N$

$$DC = 2P_{11} / (P_{1+} + P_{+1}) \quad (6)$$

$$OE = P_{21} / P_{+1} \quad (7)$$

$$CE = P_{12} / P_{1+} \quad (8)$$

## 4.6. Burned area mapping

The burned area mapping algorithm identifies changes in C-band backscatter and surface reflectance associated with burning events. The algorithm considers i) multi-temporal changes of incoherent SAR-based metrics (VV and VH backscatter coefficients and their ratio VH/VV) and ii) changes in surface reflectance (individual bands and derived indices). For algorithm training, ancillary information on thermal anomalies (hotspots) and land cover are also used.


### 4.6.1. Convolutional Neural Networks (CNN) - background

Deep learning methods are increasingly applied to remote sensing problems (Zhu et al. 2017) with convolutional neural networks (CNN) being widely used in land cover classification, the retrieval of bio-geophysical variables (Ma et al. 2019) or burned area detection and classification (Ban et al. 2020; Pinto et al. 2020). CNN are often structured by more than two stacked stages of convolution, non-linearity and pooling, followed by at least one fully connected layer (LeCun et al. 2015; Zhu et al. 2017). Each convolutional layer carries out a spatial-spectral feature extraction (Zhong et al. 2019), generating a set of new filtered data where visual and signal patterns such as edges are emphasized (Strigl et al. 2010). From the convoluted filtered data, each neuron takes a vector and applies an activation function of a weighted linear summation (9) (Maggiori et al. 2016).

$$a = f(wx + b) \quad (9)$$

where:  $a$  is the neuron output,  $w$  is the weight given to the vector  $x$ ,  $b$  is the bias value, and  $f$  is the activation function which introduces non-linearity into the network and permits learning complex features from data (Agostinelli et al. 2014; Saha et al. 2019). The most common activation function in remote sensing applications is the rectified linear unit (ReLU) (Nair and Hinton 2010), which activates values greater or equal than zero, while it converts the rest to zero (10).

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (10)$$

	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1			
		Issue	1.1	Date	27/04/2020	
				Page	17	

A loss function is used to quantify the errors when classifying a training vector data, comparing the CNN-based prediction with the label of such vector (Maggiori et al. 2016). The weights and biases of each neuron are adjusted using backpropagation during the network training, carrying out multiple iterations forward and backward (Anantrasirichai et al. 2019) to minimize the errors via gradient descent (Schmidhuber 2015). The activated data is sub-sampled often to reduce the tensor size, which increases the receptor field to the next convolutional layer of the network (Kellenberger et al. 2018; Strigl et al. 2010). The last layer of the network oversees the classification instead of the feature extraction. Thus, a fully connected neural network layer is used. Usually, such a fully connected network is followed by a softmax layer, which models the input data to the probability of belonging to each considered class (Anantrasirichai et al. 2019; Zhang et al. 2018).

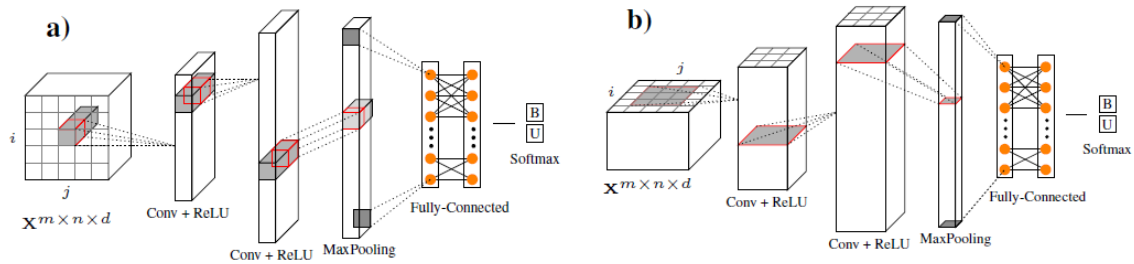
#### 4.6.2. Selection of training data

CNN is a supervised learning method and thus it needs sample data (burned and unburned pixels) for algorithm training. The extraction of the training dataset takes advantage of hotspots (active fire events) derived from thermal sensors (i.e., VIIRS and MODIS) which provide both algorithm autonomy and results replicability (i.e., avoids operator interpretation of burned/unburned areas). The use of hotspots, well established for burned area mapping (Belenguer-Plomer et al. 2019b; Roteta et al. 2019), is essential particularly when using the radar derived metrics to differentiate changes due to fires (Huang and Siegert 2006). Burned training pixels were selected within a spatial buffer determined as the double of the thermal sensor spatial resolution (Langner et al. 2007; Sitanggang et al. 2013). The unburned training pixels were those outside the hotspot buffer areas as well as from not burnable (e.g. water) land cover classes according to CCI land cover map reference.

#### 4.6.3. Assessment of optimum CNN configuration for BA mapping

The parameters that define a CNN model such as the number of layers, neurons and filters for each dataset need to be adjusted (Bashiri and Geranmayeh 2011). To determine the optimal network for BA detection and mapping, eight combinations (Table 6) were analysed as a function of network complexity (i.e., number of hidden layers), dimensionality of feature extraction (i.e., spatial or spectral) and data normalization (i.e., 0-1 or z-score). The generated CNNs were based on AlexNet (Krizhevsky et al. 2012), integrating convolutional hidden layers, the ReLU activation function, max-pooling, fully connected layers, dropout and softmax classification. Four architectures were analysed after combining two CNN-groups that differed in terms of (i) the number of hidden layers and filters, and (ii) the image domain (i.e., spatial or spectral) where the convolutional feature extraction was executed over the input data.

The first group included two CNN models with a different number of hidden layers and filters. The first model had two hidden layers with 32 and 64 filters, respectively, whereas the second model had a third additional hidden layer where 128 filters were applied. Hereafter the models with two and three hidden layers are referred to as the simple (S) and the complex (C), respectively. The second group involved two convolution-based filters for feature extraction. The first filter implies convolution over the spectral domain, considering one dimension as defined by the number of image-bands (1D). The second filter used the spatial domain at each image-band which implies two dimensions (2D) through the rows and columns (Kussul et al. 2017; Xu et al. 2017) (see Figure 5).



**Figure 5: Configuration of 1D (left) and 2D (right) CNNs (Belenguer-Plomer et al., in preparation)**

Two normalization methods were tested separately with each image band being normalized (i) in the interval [0,1] (Benedetti et al. 2018) and (ii) applying the z-score normalization (Zhong et al. 2017) (11).

$$z\text{-score}(x) = \frac{x - \mu(b)}{\sigma(b)} \quad (11)$$

where:  $x$  is a given pixel of a band  $b$  of the image, and  $\mu$  and  $\sigma$  are the mean and standard deviation, respectively. Thus, in Table 6 there are the eight configurations considered whose performance when mapping BA is going to be assessed in this study for each dataset (i.e., SAR, optical and SAR-optical combination).


**Table 6: The eight configurations assessed for each input dataset (S – simple, C – complex).**

<i>CNN model</i>	<i>Convolution dimension</i>	<i>Data normalization</i>
S	1D	z-score
S	1D	0-1
S	2D	z-score
S	2D	0-1
C	1D	z-score
C	1D	0-1
C	2D	z-score
C	2D	0-1

The utility of land cover specific CNN training was also considered as it may affect burned area mapping, particularly when using radar datasets (Belenguer-Plomer et al. 2019a; Belenguer-Plomer et al. 2018; Boschetti et al. 2004; Padilla et al. 2015).

#### 4.7. Preliminary results and significance for BA mapping algorithm development

The next sub-sections provide an overview of the BA algorithm configuration as a function of the analysis carried out to determine the optimum CNN configuration as well as the integration of SAR and optical data within a seamless mapping algorithm. Three CNN algorithms were trained and analysed: SAR based (S-1), optical based (S-2) and SAR and optical (S-1 + S-2). The latter algorithm was based on feeding both radar- and optical-derived metrics (e.g. backscatter coefficient, surface reflectance, indices) into the CNN training.

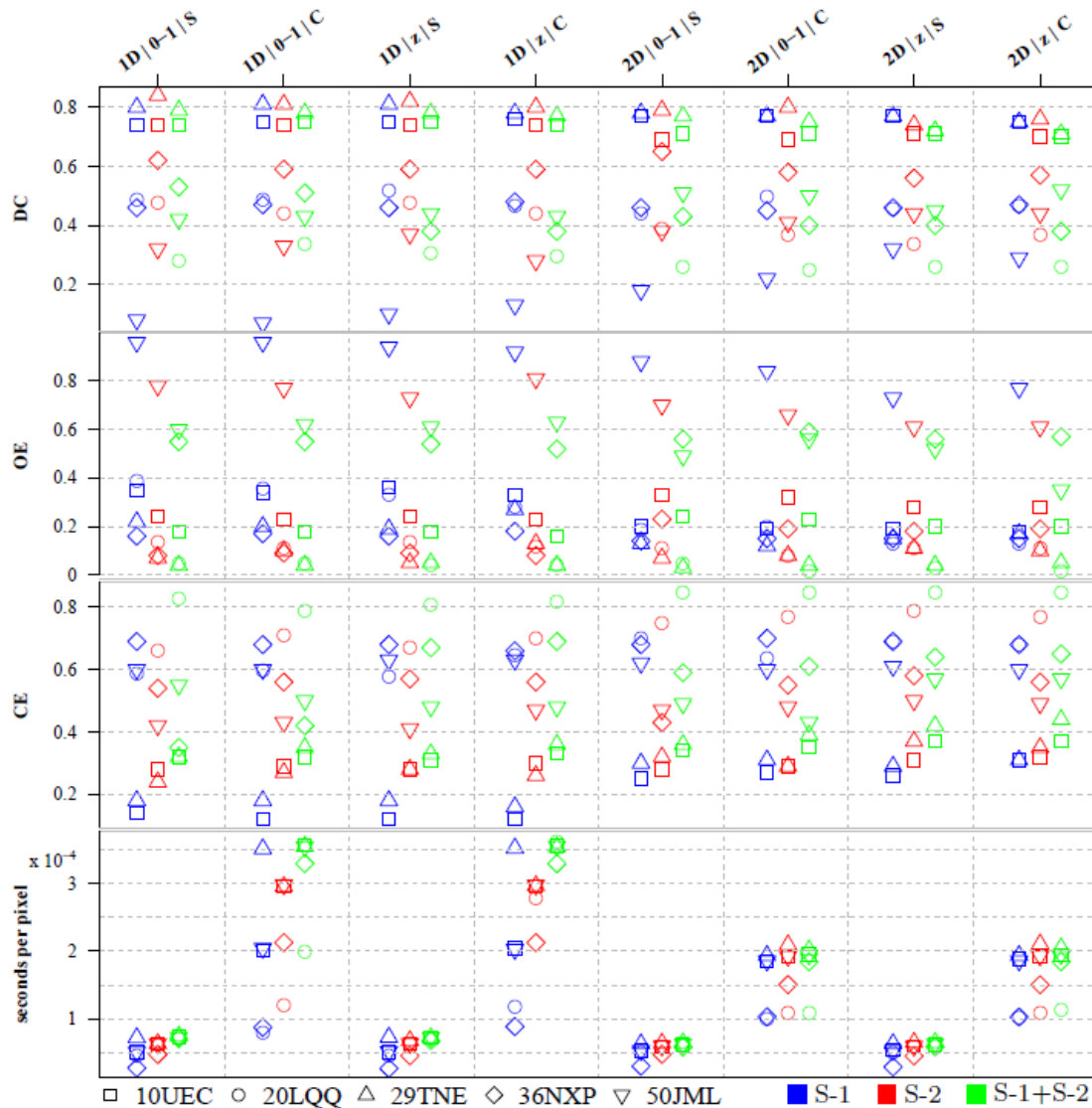
	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1		
		Issue	1.1	Date	27/04/2020
				Page	19

#### 4.7.1. Optimum CNN configuration

Depending on the MGRS tile, the optimum CNN configuration varied (Figure 6). When Sentinel-1 (S1) data was fed into the CNN, accuracy metrics of dispersion (i.e., between tiles) at any CNN configuration were larger than when compared to feeding Sentinel-2 (S2) data or both Sentinel-1 and Sentinel-2 data (S1+S2). For the radar-fed CNN (S1) inter-tiles accuracy dispersion was reduced when the convolution and feature extraction was carried out through the spatial domain of the image (2D) by decreasing omission errors (36NXP, 20LQQ and 50JML) despite a slight increase in commission errors for some tiles (10UEC and 29TNE). Similar results were achieved when using only Sentinel-2 data. When feeding both types of data (S1+S2) into the CNN, the convolution dimension (i.e., 1D, 2D) did not influence algorithm accuracy. In addition, the time required when training 2D models was lower compared to 1D, particularly when considering complex (C) networks and regardless of data normalization type. The use of more complex (C) CNN models, as opposed to using the simplest ones (S), did not result in increased accuracy, regardless of the type of data fed into the algorithm. Similarly, computing time differences as a result of different data normalization method (z-score vs 0 to 1 values) were marginal for any of the input data. However, marginal increments in accuracy were observed when using the z-score normalization for the Sentinel-1 fed algorithm, particularly in tile 50JML (Australian grasslands) where OE were reduce significantly (for 2D CNN). Conversely, when feeding Sentinel-2 or Sentinel-1 and -2 data, the 0 to 1 normalization provided slightly more accurate BA detections.

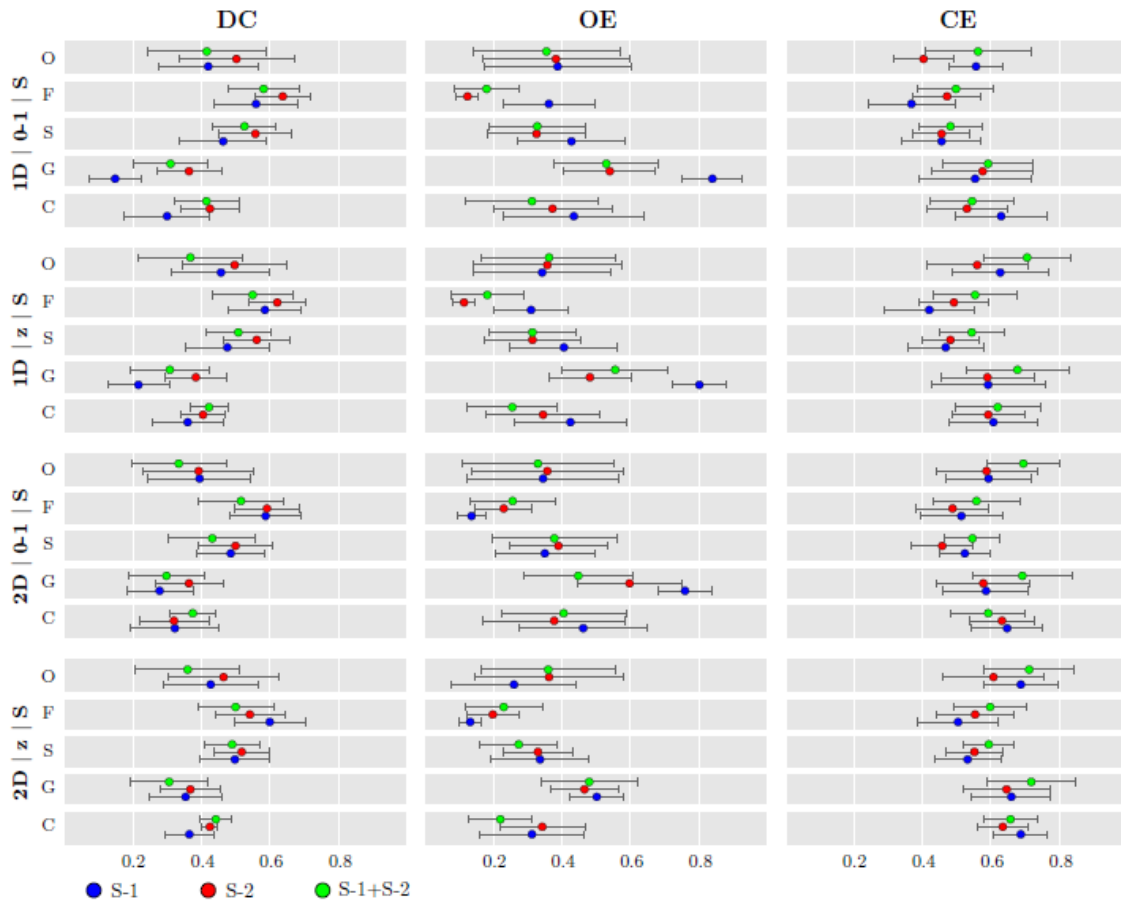


fire  
cci



**Figure 6: Dice coefficient (DC), commission and omission errors (CE and OE) by calibration tiles when training different CNN configuration and input data.**

The results show that more complex networks are not adequate when mapping BA from either SAR, optical or both datasets since training is more computationally intensive but BA accuracy does not improve. However, convolution dimensionality and data normalization warrant further investigation as their effect changes with the studied area, i.e., vegetation type and fire regimes (Figure 7).



**Figure 7: Mean and standard error of Dice coefficient (DC), commission and omission errors (CE and OE) by land cover classes when training different CNN configuration and input data. O: other. F: forest. S: shrubland. G: grassland. C: cropland.**

By land cover classes, the lowest BA mapping accuracy was observed for Grasslands, particularly when using Sentinel-1 data due to high omission errors. However, combining 2D convolution with z-score normalization resulted in improving the DC by 59% from 1D convolution-based approaches with z-score (DC  $0.35 \pm 0.24$  vs  $0.22 \pm 0.2$ , being  $\pm$  the standard deviation). The same configuration (2D and z-score) also improved the accuracy over Crops, especially when compared to 1D with 0-1 data normalization (DC  $0.37 \pm 0.14$  vs  $0.30 \pm 0.25$ ), although to a lesser extent, while over Forests the improvement was marginal. Accuracy metrics were stable for Shrubs over all configurations tested, although the 2D and z-score configuration provided less overall dispersion among the analysed tiles. For the remaining land cover types, included in class Others, the highest mapping accuracy was achieved using the convolution in spectral domain (1D).

Although S-2 fed CNN achieved better accuracies when compared to S-1 fed CNN, such improvements were also conditioned by land cover classes and configurations. When using optical data, the most accurate dimension for convolution was the spectral one (1D) except for Crops where the spatial convolution (2D) provided improved results. Marginal differences in BA accuracy were found between the two data normalization types with the z-score normalization providing higher DC values over all land cover classes except for Forests.



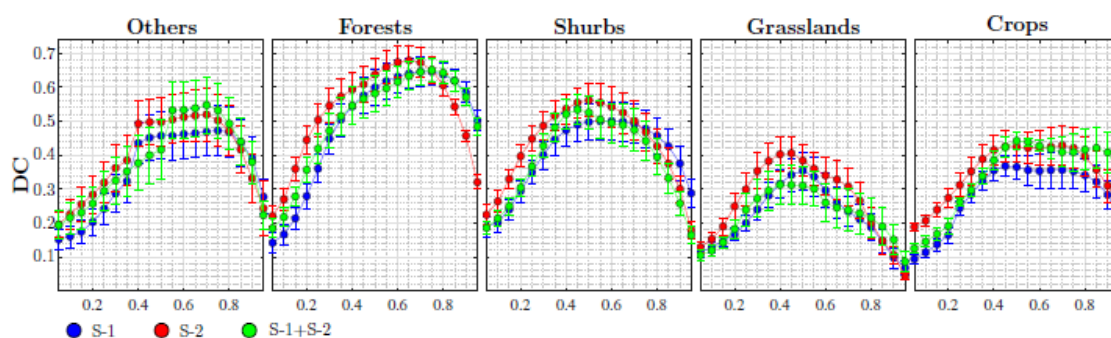
When both, Sentinel-1 and Sentinel-2 data were fed to the CNN algorithm the BA classification did not improved when compared to only using Sentinel-2 data except for Crops. Over cropping areas, SAR or optical data alone provided low mapping accuracies (highest DCs achieved  $0.37\pm 0.14$  and  $0.42\pm 0.05$ , respectively). However, combining the two sensor types (S1+S2) improved the mapping accuracy (DC  $0.44\pm 0.09$ ) by reducing OE. Such an improvement was maximum for the 2D convolution and z-score normalization. For the remaining land cover classes, mixing SAR and optical data seem not necessary when cloud cover is not an issue. As for the CNN optimum configuration, 1D convolution and 0 to 1 normalization provided improved mapping accuracies as for S-1 based algorithm.

The highest mapping accuracy was observed for Forests regardless of data normalization method, convolution dimensionality and input remote sensing data (S-1, S-2, S1+S2). The optimum CNN configuration for each land cover class is presented in Table 7 as a function of the input remote sensing data.

**Table 7: Optimum CNN configuration and the average Dice Coefficient of the calibration tiles (DC  $\pm$  standard deviation) achieved by land cover class and input data**

LC	S-1	DC (S-1)	S-2	DC (S-2)	S-1+S-2	DC (S-1+S-2)
O	1D   z-score	$0.46\pm 0.29$	1D   z-score	$0.50\pm 0.31$	1D   0-1	$0.42\pm 0.38$
F	2D   z-score	$0.60\pm 0.23$	1D   0-1	$0.64\pm 0.18$	1D   0-1	$0.58\pm 0.24$
S	2D   z-score	$0.50\pm 0.23$	1D   z-score	$0.56\pm 0.22$	1D   0-1	$0.53\pm 0.21$
G	2D   z-score	$0.35\pm 0.24$	1D   z-score	$0.38\pm 0.20$	all	$0.31\pm 0.25$
C	2D   z-score	$0.37\pm 0.14$	2D   z-score	$0.43\pm 0.05$	2D   z-score	$0.44\pm 0.09$


The layer of the CNNs (softmax) predicts the probability of each pixel to belong to burned and unburned classes, and thus results into mapped burned and unburned categories. All previous analyses assigned a pixel as burned when the softmax probability was above 50%. However, such a fixed threshold may not always provide optimum results depending on the input datasets and land cover class (Figure 8).



**Figure 8: Variation of mapping accuracy as a function of changes in softmax probability by land cover class and input data sets**

To improve BA mapping accuracy and balanced CE and OE the effect of a using a variable threshold probability was studied. Such variation depended on land cover class and the input data fed to the CNN algorithm. Over Grasslands, Crops and Shurbs, the classes with the highest OE, improved accuracies where observed when the softmax



	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1		
		Issue	1.1	Date	27/04/2020
				Page	23

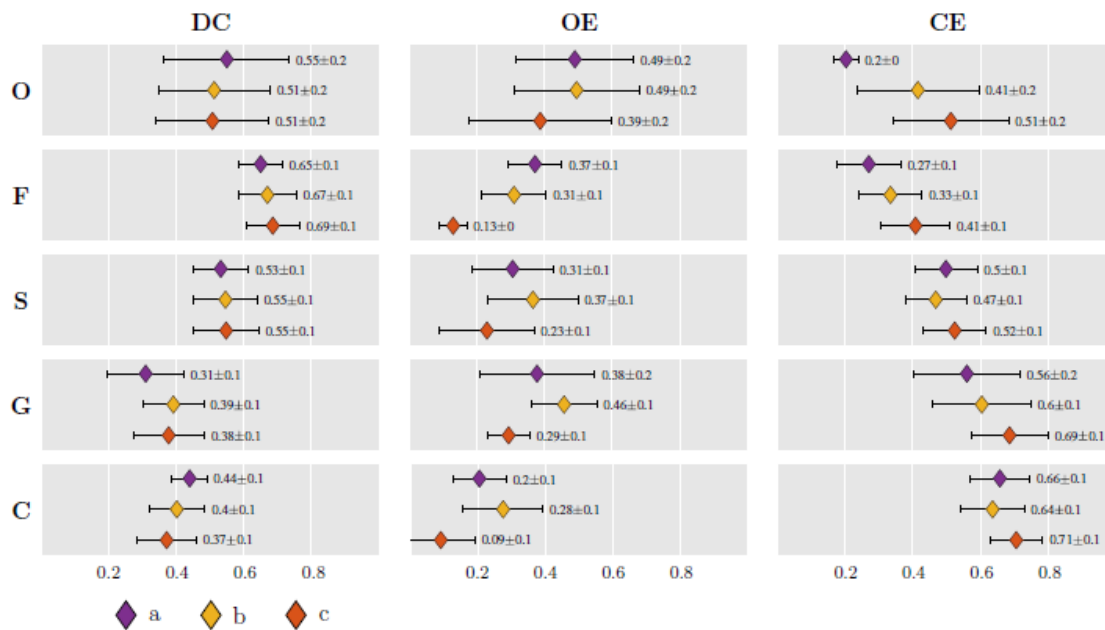
probability was reduced (40 to 50%) with its extent depending on the input dataset. The best trade-off between CE and OE was provided by thresholds of 45% and 40% over Shrubs and respectively Grasslands, when using S1+S2 data. When S1 or S2 data were used separately a 45% threshold was optimum over Crops and Grasslands. Conversely, for Forests class a more restrictive threshold improved BA mapping accuracy. The optimum threshold differed with the input data, from 65% when using S1 or S2 data alone to 75% when integrating SAR and optical data (S1+S2). For class Others the BA accuracy improved marginally when varying the threshold until a probability of 80% for S1 and 70% for S2. However, when using S1+S2 the improvement was considerable with within the 55-75% interval with the highest accuracy being achieved for a softmax of 70%. Such improvement allowed S1+S2 based detection achieving higher accuracies when compared to separate S1 or S2 detection. Past the optimum threshold, mapping accuracy reduces considerably, particularly when using S2 data. This effect was observed for all land cover classes except Grasslands, where the opposite was true.

#### 4.7.2. SAR-optical mapping strategy

BA mapping accuracy was comparatively analysed by land cover class for three different strategies of combining SAR and optical datasets: (a) stacking radar and optical metrics (e.g., backscatter coefficient, surface reflectance, indices) and feed them to the CNN algorithm; (b) using BA detected from the optical data and filling the gaps (e.g., due to cloud cover) with pixels mapped from radar data and; (c) joining BA detected by either radar or the optical-based algorithms (Figure 9).

Over Forests, the three mapping strategies provided similar results (i.e., DC values). However, joining individual S-1 and S-2 detections may provide an advantage by reducing areas not detected due to clouds or shadows. For Shrubs the observed DC values were similar for all mapping strategies with radar-filled optical-based BA maps showing slightly higher DCs when compared to the remaining two strategies. Over Grasslands the radar-filled optical-based BA maps, provided more accurate results when compared to the remaining ones.

Over the two remaining land cover classes (i.e., Others, Crops), the use of radar-optical stacks into the CNN allowed for improved accuracies when compared to using radar or optical data separately and combining the result. In particular, for class Others the radar-optical stacks allowed for reduced CE (20% less).



**Figure 9: Mean and standard error of Dice coefficient (DC), commission and omission errors (CE and OE) by land cover classes when combining Sentinel-1 and -2 applying four different approaches, being a - a SAR and optical data integration before the CNN based classification; b – the Sentinel-2 BA based maps pixels affected by no data pixels were filled using BA from Sentinel-1; and c - a SAR and optical integration of all burned pixels.**

#### 4.8. Preliminary algorithm validation

The optimum CNN configuration and mapping strategy, as observed over the calibration tiles, was assessed over the validation tiles (Table 8) with the mapping accuracy varying depending on the input data (i.e., S-1, S-2 and S-1+S-2). Higher mapping errors (DC<0.6) were observed over grassland dominated tiles in Africa and Australia (33NTG and 52LCH, respectively), regardless of the input data. Over the remaining tiles, DC values were above 0.7. Over two tiles (20LQP and 33NTG), the radar-based maps were more accurate when compared to the optical based maps (DC of 0.81 vs 0.71 and 0.50 vs 0.46) with the opposite being valid for the remaining three tiles. However, the use of Sentinel-1 data (cloud cover independent) allowed for wall to wall mapping (marginal percentages of unobserved pixels at image borders). In tile 52LCH the optical-based maps did not provide information for 21.5% when compared to 0.3% when using the radar data (Figure 10).

By land cover type (Figure 11), the highest accuracy was observed over forested areas when mapping BA through the SAR-optical combination (DC 0.72) as opposed to only using SAR (DC 0.63) or optical (DC 0.66) information (Figure 11). The most relevant improvement was combining S-1 and S-2 was found over class Others where the synergy of both sensors reduced considerably OE and CE when compared to the single sensor approaches. Over the remaining land cover classes, the radar-optical approach provided the lowest OE, particularly when compared to the SAR-based maps. The CE values of the radar-optical combination were equal or marginally higher when compared to maps derived from Sentinel-2 optical data. Finally, the lowest accuracy was achieved over the Crops, mainly due to unusually high CE (near of 0.8) observed for both sensor types when compared to the other land cover types. In addition, for the radar-based maps, BA

accuracy over cropping areas was also negatively influenced by high OE, which did not occur when using optical datasets.

The optimum combination of Sentinel-1 and Sentinel-2 data generally improved or maintained the accuracy achieved from maps based on SAR or optical data separately except for tile 20LQP, where the SAR-based maps achieved the highest. Improvements when combining the two sensor types were related to a considerable reduction in OE coupled with marginal increase in CE. The average OE reduction and CE increase over the five validation tiles was  $0.22 \pm 0.22$  and  $0.05 \pm 0.17$  when compared to radar-based maps and  $0.09 \pm 0.08$  and  $0.05 \pm 0.05$  when compared to optical-based maps. Apart from accuracy improvements, using both optical and radar data reduced gaps due to cloud cover to almost non-existent, a major advantage of combining active and passive sensors. Although computationally intensive, a combined SAR-optical approach takes advantage of the optical sensor improved BA mapping accuracy and the radar atmospheric independence to provide consistent, wall to wall products.

**Table 8: Error metrics for Sentinel-1 burned area detections for each MGRS tile analysed**

MGRS	C	Reference period	Sat	Detection period	DC	OE	CE	%Nd
10SEH	NA	04/10/2017–05/11/2017	S-1	28/09/2017–03/11/2017	0.46	0.69	0.13	0.00
			S-2	07/10/2017–01/11/2017	0.70	0.12	0.41	2.26
			S-1+S-2	28/09/2017–03/11/2017	0.70	0.10	0.43	0.00
20LQP	SA	20/07/2016–22/09/2016	S-1	03/07/2016–25/09/2016	0.81	0.08	0.27	0.00
			S-2	17/07/2016–25/09/2016	0.71	0.20	0.37	0.00
			S-1+S-2	03/07/2016–25/09/2016	0.73	0.04	0.41	0.00
29TNG	Eu	05/10/2017–06/11/2017	S-1	28/09/2017–09/11/2017	0.64	0.44	0.25	0.00
			S-2	05/10/2017–09/11/2017	0.75	0.27	0.22	0.06
			S-1+S-2	28/09/2017–09/11/2017	0.77	0.23	0.22	0.00
33NTG	Af	15/01/2016–16/02/2016	S-1	15/01/2016–20/02/2016	0.50	0.53	0.47	0.00
			S-2	18/01/2016–17/02/2016	0.47	0.65	0.31	0.39
			S-1+S-2	15/01/2016–20/02/2016	0.56	0.47	0.42	0.00
52LCH	Au	05/04/2017–21/04/2017	S-1	26/03/2017–19/04/2017	0.36	0.75	0.34	0.00
			S-2	19/03/2017–08/04/2017	0.55	0.59	0.15	17.56
			S-1+S-2	26/03/2017–19/04/2017	0.56	0.55	0.24	0.00

C - continent for each tile (Af-Africa, Au-Australia, Eu-Europe, NA-North America and SA-South America); Reference period - period for which it was derived the reference burn perimeters using Landsat-8; Sat - input dataset considered; Detection period - first and last Sentinel-1 or Sentinel-2 images of the temporal series; DC - Dice Coefficient; OE - Omission Error; CE - Commission Error; and %Nd - the percentage of no data pixels over all the MGRS tile.



fire  
cci

Fire\_cci  
SAR-Optical Algorithm  
Theoretical Basis Document

Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1		
Issue	1.1	Date	27/04/2020
		Page	26

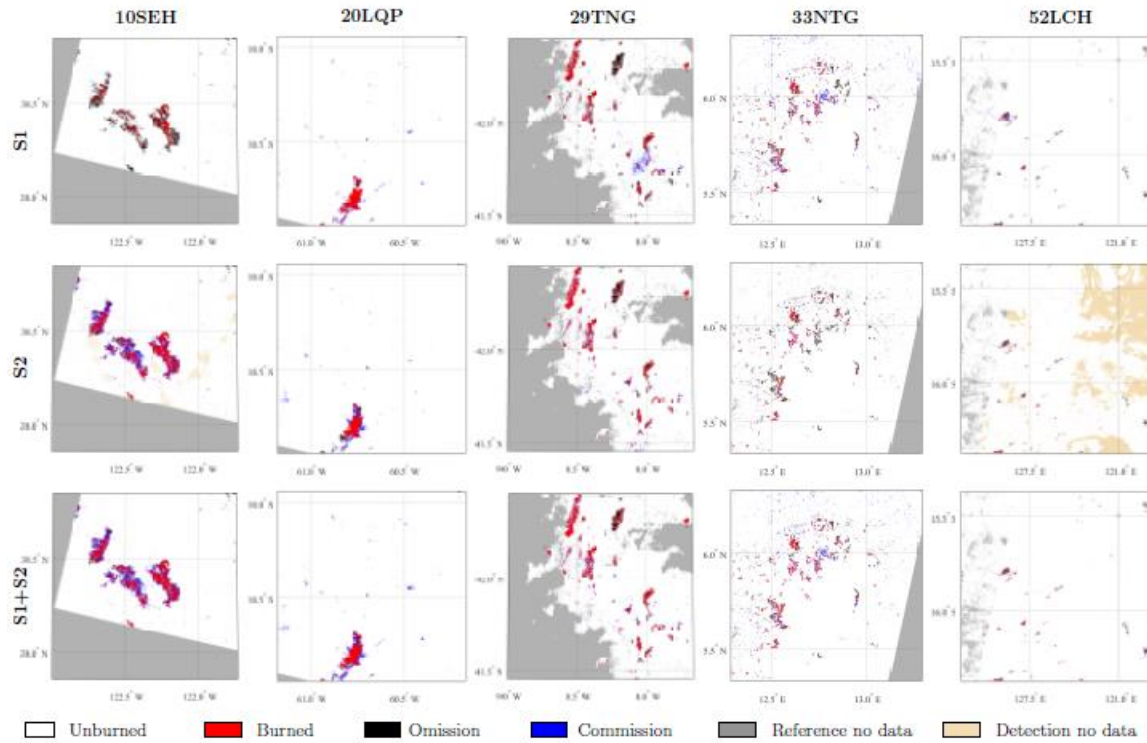


Figure 10: BA mapping from active (S1), passive (S2) and active-passive (S1+S2) sensors

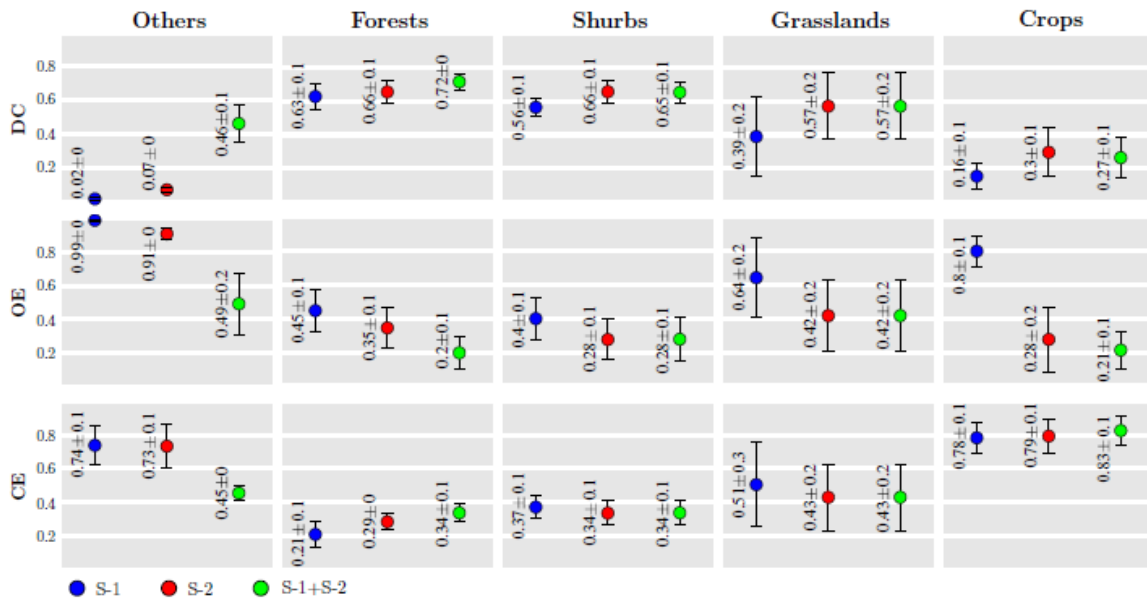




Figure 11: BA mapping accuracy for the validating tiles as a function of the sensor used

	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1		
		Issue	1.1	Date	27/04/2020
				Page	27


## 5. References

- Agostinelli, F., Hoffman, M., Sadowski, P., & Baldi, P. (2014). Learning activation functions to improve deep neural networks. *arXiv preprint arXiv:1412.6830*
- Anantrasirichai, N., Biggs, J., Albino, F., & Bull, D. (2019). A deep learning approach to detecting volcano deformation from satellite imagery using synthetic datasets. *Remote Sensing of Environment*, 230, 111179-111179
- Andreae, M.O., & Merlet, P. (2001). Emission of trace gases and aerosols from biomass burning. *Global Biogeochemical Cycles*, 15, 955-966
- Ban, Y., Zhang, P., Nascetti, A., Bevington, A.R., & Wulder, M.A. (2020). Near Real-Time Wildfire Progression Monitoring with Sentinel-1 SAR Time Series and Deep Learning. *Scientific Reports*, 10, 1–15-11–15
- Bashiri, M., & Geranmayeh, A.F. (2011). Tuning the parameters of an artificial neural network using central composite design and genetic algorithm. *Scientia Iranica*, 18, 1600–1608-1600–1608
- Belenguier-Plomer, M.A., Chuvieco, E., & Tanase, M.A. (2019a). *Evaluation of backscatter coefficient temporal indices for burned area mapping: SPIE*
- Belenguier-Plomer, M.A., Tanase, M.A., Fernandez-Carrillo, A., & Chuvieco, E. (2018). *Insights into burned areas detection from Sentinel-1 data and locally adaptive algorithms: SPIE*
- Belenguier-Plomer, M.A., Tanase, M.A., Fernandez-Carrillo, A., & Chuvieco, E. (2019b). Burned area detection and mapping using Sentinel-1 backscatter coefficient and thermal anomalies. *Remote Sensing of Environment*, 233, 111345
- Benedetti, P., Ienco, D., Gaetano, R., Ose, K., Pensa, R.G., & Dupuy, S. (2018). M<sup>3</sup> Fusion: A Deep Learning Architecture for Multiscale Multimodal Multitemporal Satellite Data Fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11, 4939–4949-4939–4949
- Boschetti, L., Flasse, S.P., & Brivio, P.A. (2004). Analysis of the conflict between omission and commission in low spatial resolution dichotomic thematic products: The Pareto Boundary. *Remote Sensing of Environment*, 91, 280–292-280–292
- Bourgeau-Chavez, L.L., Kasischke, E.S., Brunzell, S., & Mudd, J.P. (2002). Mapping fire scars in global boreal forests using imaging radar data. *International Journal of Remote Sensing*, 23, 4211-4234
- Chuvieco, E. (1999). Measuring changes in landscape pattern from satellite images: short-term effects of fire on spatial diversity. *International Journal of Remote Sensing*, 20, 2331-2346
- Chuvieco, E., Yue, C., Heil, A., Mouillot, F., Alonso-Canas, I., Padilla, M., Pereira, J.M., Oom, D., & Tansey, K. (2016). A new global burned area product for climate assessment of fire impacts. *Global Ecology and Biogeography*, 25, 619-629
- ESA. Land Cover CCI Product User Guide Version 2. Tech. Rep. (2017). Available at: [http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2\\_2.0.pdf](http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf).




	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1			
		Issue	1.1	Date	27/04/2020	
				Page	28	

- Giglio, L., Loboda, T., Roy, D.P., Quayle, B., & Justice, C.O. (2009). An active-fire based burned area mapping algorithm for the MODIS sensor. *Remote Sensing of Environment*, *113*, 408-420
- Giglio, L., Randerson, J.T., & Werf, G.R. (2013). Analysis of daily, monthly, and annual burned area using the fourth generation global fire emissions database (GFED). *Journal of Geophysical Research: Biogeosciences*, *118*, 317-328
- Huang, S., & Siegert, F. (2006). Backscatter change on fire scars in Siberian boreal forests in ENVISAT ASAR wide-swath images. *IEEE Geoscience and Remote Sensing Letters*, *3*, 154–158-154–158
- Kellenberger, B., Marcos, D., & Tuia, D. (2018). Detecting mammals in UAV images: Best practices to address a substantially imbalanced dataset with deep learning. *Remote Sensing of Environment*, *216*, 139–153-139–153
- Key, C.H., & Benson, N.C. (2006). Landscape assessment (LA) Sampling and Analysis Methods. In D.C. Lutes, R.E. Keane, J.F. Caratti, C.H. Key, N.C. Benson, S. Sutherland & L.J. Gangi (Eds.), *FIREMON: Fire effects monitoring and inventory system* (pp. 1-55). Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Gen. Tech. Rep. RMRS-GTR-164-CD
- Kloster, S., Mahowald, N.M., Randerson, J.T., & Lawrence, P.J. (2012). The impacts of climate, land use, and demography on fires during the 21st century simulated by CLM-CN. *Biogeosciences*, *9*
- Krizhevsky, A., Sutskever, I., & Hinton, G.E. (2012). Imagenet classification with deep convolutional neural networks. In, *Advances in neural information processing systems* (pp. 1097–1105)
- Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, *14*, 778–782-778–782
- Laneve, G., Castronuovo, M.M., & Cadau, E.G. (2006). Continuous Monitoring of Forest Fires in the Mediterranean Area Using MSG. *IEEE Transactions on Geoscience and Remote Sensing*, *44*, 2761-2768
- Langner, A., Miettinen, J., & Siegert, F. (2007). Land cover change 2002–2005 in Borneo and the role of fire derived from MODIS imagery. *Global Change Biology*, *13*, 2329–2340-2329–2340
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*, 436-436
- Lohberger, S., St€angel, M., Atwood, E.C., & Siegert, F. (2017). Spatial evaluation of Indonesia’s 2015 fire-affected area and estimated carbon emissions using Sentinel-1. *Global Change Biology*
- López-García, M.J., & Caselles, V. (1991). Mapping burns and natural reforestation using Thematic Mapper data. *Geocarta International*, *61*, 31-37
- Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B.A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *152*, 166–177-166–177

	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1			
		Issue	1.1	Date	27/04/2020	
					Page	29

- Maggiore, E., Tarabalka, Y., Charpiat, G., & Alliez, P. (2016). Convolutional neural networks for large-scale remote-sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *55*, 645–657–645–657
- Mouillot, F., Schultz, M.G., Yue, C., Cadule, P., Tansey, K., Ciais, P., & Chuvieco, E. (2014). Ten years of global burned area products from spaceborne remote sensing— A review: Analysis of user needs and recommendations for future developments. *International Journal of Applied Earth Observation and Geoinformation*, *26*, 64-79
- Nair, V., & Hinton, G.E. (2010). Rectified linear units improve restricted boltzmann machines. In, *Proceedings of the 27th international conference on machine learning (ICML-10)* (pp. 807–814-807–814)
- Padilla, M., Stehman, S.V., & Chuvieco, E. (2014). Validation of the 2008 MODIS-MCD45 global burned area product using stratified random sampling. *Remote Sensing of Environment*, *144*, 187-196
- Padilla, M., Stehman, S.V., Hantson, S., Oliva, P., Alonso-Canas, I., Bradley, A., Tansey, K., Mota, B., Pereira, J.M., & Chuvieco, E. (2015). Comparing the Accuracies of Remote Sensing Global Burned Area Products using Stratified Random Sampling and Estimation. *Remote Sensing of Environment*, *160*, 114-121
- Padilla, M., Wheeler, J., & Tansey, K. (2017). ESA CCI ECV Fire Disturbance: D4.1.1. Product Validation Report, version 1.0. In A.a. <http://www.esa-fire-cci.org/documents> (Ed.)
- Pinto, M.M., Libonati, R., Trigo, R.M., Trigo, I.F., & DaCamara, C.C. (2020). A deep learning approach for mapping and dating burned areas using temporal sequences of satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *160*, 260–274-260–274
- Randerson, J.T., Chen, Y., Werf, G.R., Rogers, B.M., & Morton, D.C. (2012). Global burned area and biomass burning emissions from small fires. *Journal of Geophysical Research: Biogeosciences*, *117*
- Rignot, E.J.M., & Zyl, J.J.v. (1993). Change Detection Techniques for ERS-1 SAR Data. *IEEE Transactions on Geoscience and Remote Sensing*, *31*, 896-906
- Roteta, E., Bastarrika, A., Padilla, M., Storm, T., & Chuvieco, E. (2019). Development of a Sentinel-2 burned area algorithm: Generation of a small fire database for sub-Saharan Africa. *Remote Sensing of Environment*, *222*, 1-17
- Saha, S., Bovolo, F., & Bruzzone, L. (2019). Unsupervised Deep Change Vector Analysis for Multiple-Change Detection in VHR Images. *IEEE Transactions on Geoscience and Remote Sensing*, *57*, 3677-3693
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, *61*, 85–117-185–117
- Simmonds, P.G., Manning, A.J., Derwent, R.G., Ciais, P., Ramonet, M., Kazan, V., & Ryall, D. (2005). A burning question. Can recent growth rate anomalies in the greenhouse gases be attributed to large-scale biomass burning events? *Atmospheric Environment*, *39*, 2513-2517



	Fire_cci SAR-Optical Algorithm Theoretical Basis Document	Ref.:	Fire_cci_D2.1.1_ATBD_SAR-O_1.1			
		Issue	1.1	Date	27/04/2020	
				Page	30	

- Sitanggang, I.S., Yaakob, R., Mustapha, N., & Ainuddin, A.N. (2013). Predictive models for hotspots occurrence using decision tree algorithms and logistic regression. *Journal of applied sciences*, *13*, 252–261-252–261
- Strigl, D., Kofler, K., & Podlipnig, S. (2010). Performance and scalability of GPU-based convolutional neural networks. In, *18th Euromicro Conference on Parallel, Distributed and Network-based Processing* (pp. 317–324)
- Stroppiana, D., Tansey, K., Grégoire, J.-M., & Pereira, J.M.C. (2003). An Algorithm for Mapping Burnt Areas in Australia Using SPOT-VEGETATION Data. *IEEE Transactions on Geoscience and Remote Sensing*, *41*, 907-909
- Tanase, M.A., de la Riva, J., Santoro, M., Le Toan, T., & Perez-Cabello, F. (2010). Sensitivity of X-, C- and L-band SAR backscatter to fire severity in mediterranean pine forests. *IEEE Transactions on Geoscience and Remote Sensing*, *48*, 3663-3675
- van der Werf, G.R., Randerson, J.T., Giglio, L., van Leeuwen, T.T., Chen, Y., Rogers, B.M., Mu, M., van Marle, M.J.E., Morton, D.C., Collatz, G.J., Yokelson, R.J., & Kasibhatla, P.S. (2017). Global fire emissions estimates during 1997–2016. *Earth Syst. Sci. Data*, *9*, 697-720
- Xu, X., Li, W., Ran, Q., Du, Q., Gao, L., & Zhang, B. (2017). Multisource remote sensing data classification based on convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, *56*, 937–949-937–949
- Zhang, C., Pan, X., Li, H., Gardiner, A., Sargent, I., Hare, J., & Atkinson, P.M. (2018). A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *140*, 133–144-133–144
- Zhong, L., Hu, L., & Zhou, H. (2019). Deep learning based multi-temporal crop classification. *Remote Sensing of Environment*, *221*, 430–443-430–443
- Zhong, Y., Fei, F., Liu, Y., Zhao, B., Jiao, H., & Zhang, L. (2017). SatCNN: satellite image dataset classification using agile convolutional neural networks. *Remote Sensing Letters*, *8*, 136–145-136–145
- Zhu, X.X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, *5*, 8–36-38–36



## Annex 1: Acronyms and abbreviations

AOI	Area Of Interest
ATBD	Algorithm Theoretical Basis Document
BA	Burned Area
BOA	Bottom of atmosphere
CCI	Climate Change Initiative
CE	Commission Error
CESBI O	Centre for the Study of the Biosphere from Space
CNES	France National Centre for Space Studies
CNN	Convolution neural networks
DC	Dice Coefficient
DEM	Digital Elevation Model
DIAS	Data and Information Access Services
DL	Deep Learning
ECV	Essential Climate Variables
ESA	European Space Agency
EW	Extra Wide swath mode
FIRMS	Fire Information for Resource Management System
GCOS	Global Climate Observing System
GRD	Ground Range Detected
IPCC	Intergovernmental Panel on Climate Change
IW	Interferometric Wide swath mode
LC	Land Cover
MGRS	Military Grid Reference System
MIRBI	Mid-InfraRed Burnt Index
MODIS	Moderate-Resolution Imaging Spectroradiometer
MSI	Multi Spectral Instrument
NASA	National Aeronautics and Space Agency

NBR	Normalized Burn Ratio
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Near InfraRed
OE	Omission Error
OLI	Operational Land Imager
OTB	Orfeo ToolBox
PEPS	Plateforme d'Exploitation des Produits Sentinel
R	Red
RF	Random Forest
RGB	Red-Green-Blue composite
ReLU	Rectified Linear Unit
RXD	Reed-Xiaoli Detector
S1	Sentinel-1
S2	Sentinel-2
SAR	Synthetic Aperture Radar
SAR-O	SAR-Optical
SCL	Scene Classification
SFD	Small Fire Database
SLC	Single Look Complex
SM	Strip Map mode
SRTM	Shuttle Radar Topography Mission
SWIR	Short Wave InfraRed
TIRS	Thermal InfraRed Sensor
TOPSA R	Terrain Observation with Progressive Scans SAR
USGS	United States Geological Survey
UTM	Universal Transverse Mercator
VIIRS	Visible Infrared Imaging Radiometer Suite
WGS84	World Geodetic System 1984