ESA Climate Change Initiative – Fire_cci
D2.2 End to End ECV Uncertainty Budget (E3UB)

Project Name	ECV Fire Disturbance: Fire_cci
Contract Nº	4000126706/19/I-NB
Issue Date	16/08/2021
Version	2.1
Author	Amin Khaïroun, Joshua Lizundia-Loiola, Gonzalo Otón, Mihai A. Tanase, M. Lucrecia Pettinari, Emilio Chuvieco
Document Ref.	Fire_cci_D2.2_E3UB_v2.1
Document type	Public

Project Partners

Prime Contractor/ Scientific Lead & Project Management	UAH – University of Alcala (Spain)
Earth Observation Team	UAH – University of Alcala (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 - Laboratory for Sciences of Climate and Environment (France)

Distribution

Affiliation	Name	Address	Copies
ESA	Clément Albergel (ESA)	clement.albergel@esa.int	electronic copy
Project Team	Emilio Chuvieco (UAH)	emilio.chuvieco@uah.es	electronic copy
	M. Lucrecia Pettinari (UAH)	mlucrecia.pettinari@uah.es	
	Amin Khaïroun (UAH)	amin.khairoun@uah.es	
	Joshua Lizundia (UAH)	joshua.lizundia@uah.es	
	Gonzalo Otón (UAH)	gonzalo.oton@uah.es	
	Mihai Tanase (UAH)	mihai.tanase@uah.es	
	Consuelo Gonzalo (UPM)	consuelo.gonzalo@upm.es	
	Dionisio Rodríguez Esparragón (UPM)	dionisio.rodriguez@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	
	Idir Bouarar (MPIM)	idir.bouarar@mpimet.mpg.de	
	Florent Mouillot (CNRS)	florent.mouillot@cefe.cnrs.fr	
	Philippe Ciais (CNRS)	philippe.ciais@lsce.ipsl.fr	

Summary

This End-to-End ECV Uncertainty Budget (E3UB) document describes the analyses that were made to identify and estimate the major sources of error that arise in each step of the Fire Disturbance ECV's final product retrieval process.

	Affiliation/Function	Name	Date
Prepared	UAH	Amin Khaïroun Joshua Lizundia-Loiola Gonzalo Otón Mihai A. Tanase	16/08/2021
Reviewed	UAH – Project Manager	Lucrecia Pettinari	16/08/2021
Authorized	UAH - Science Leader	Emilio Chuvieco	16/08/2021
Accepted	ESA - Technical Officer	Clément Albergel	31/08/2021

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

Document Status Sheet

Issue	Date	Details
1.0	30/03/2020	First Version of the document
1.1	30/04/2020	Addressing comments of Fire_cci+_D2.2_E3UB_v1.0_RID.doc
2.0	23/07/2021	Second version of the document
2.1	16/08/2021	Addressing comments of Fire_cci+_D2.2_E3UB_v2.0_RID.doc

Document Change Record

Issue	Date	Request	Location	Details
1.1	30/04/2020	ESA	Section 2 Section 4 and sub-sections Section 4.3 Section 6	Small changes in the text Sub-sections reorganized, with the ones corresponding to FireCCILT20 and FireCCIS1S2AF10 merged and expanded. Text expanded. Additional references added.
2.0	23/07/2021	UAH	Sections 1, 3.1, 5.1, Annex 1 Sections 2.1, 4.1, 4.2, 4.3, 4.4, 4.5, 5, 6 Sections 2.2, 2.3, 2.4, 5.2	Small changes in the text Text updated New sections added
2.1	16/08/2021	ESA	All sections Section 2.4 Section 3.3 Sction 4.2 Section 4.3	Added numbers to the equations. Paragraph added at the end of the section. Added link to SYN ATBD. Caption of Figure 5 updated. First paragraph rephrased.

Table of Contents

List of Tables

List of Figures

1 Executive Summary

This version of the End-to-End ECV Uncertainty Budget (E3UB) document aims to identify, characterise and estimate the major sources of error that arise in each step of the ECV's final product retrieval process, as well as, to update the methods introduced in the previous version (E3UB version 1.1), with more convenient approaches of uncertainty propagation for the different products. In general terms, each step is related to a processing level. Those levels are typically known as Level-0, Level-1, Level-2, Level-3 (gridded data) and Level-4, although this naming convention has not been officially adopted by every space agency yet. This document tries to determine and estimate the distribution (uncertainty) that is followed by the errors that are caused by different effects throughout those levels, and provides a frame for their propagation. To do that, input sensors that were used by the different algorithms developed within the project for the burned area retrieval (Sentinel-1 SAR, Sentinel-2 MSI, Sentinel-3 Synergy, NOAA AVHRR, and Terra MODIS), and their corresponding ancillary data (Land Cover CCI, Active Fires, etc.) were properly characterised.

2 Concepts and methods for Uncertainty characterisation

2.1 Introduction

Uncertainty analysis and validation are critical phases to generate any Essential Climate Variable (ECV), and therefore both have an important role in the European Space Agency's (ESA) Climate Change Initiative (CCI) programme since its inception. The Guide to the Expression of Uncertainty in Measurement (GUM) that is maintained by the Joint Committee for Guides in Metrology (JCGM) describes this process as the "general law of error propagation", and defines the uncertainty of measurement as "a parameter, associated with the result of a measurement that characterizes the dispersion of the values that could reasonably be attributed to the measurand" (GUM, 2008). While the validation gives information about the global quality of the product, the uncertainty analysis tries to find out which factors are affecting those results and how they are distributed. This implies characterising every source of error and determining how they are propagated through the entire retrieval process, i.e. from the input data to the product estimates through the functions used to derive the estimates.

European research projects such as QA4EO and FIDUCEO have developed guidelines to concretise the application of these generic uncertainty propagation methods to Earth Observation (EO) data at a pixel level (Mittaz et al., 2019). However, it is important to separate two different stages of uncertainty characterisation that are clearly discernible in any EO data application. The first one encompassed the uncertainties of all processing levels that were applied prior to the usage of the EO data by the algorithm developer. In the case of burned area (BA) estimation, this refers to all the processes applied from raw data to generate the Level-1 (radiances) and Level-2 (reflectance) products. These processes occur before applying the BA algorithms, and therefore should be the responsibility of those in charge of the pre-processing algorithms. The second stage of uncertainty characterization deals with those factors related to the BA algorithm, which could be considered as the actual responsibility of the BA algorithm developers. They include all processes and transformations required to classify pixels as burned or not, as well as derived products (burned area estimates, for instance), which are offered to a higher-level user. In the case of Fire_cci BA, this stage refers to the uncertainties associated with BA algorithms or any process used to produce the final pixel and grid products (see definition of both products in the Product Specification Document (PSD) of Fire_cci Phase

2 (Chuvieco et al., 2017) and its update in the Algorithm Development Plan (ADP) document (Pettinari et al., 2019).

The definition of the previous two phases implies that the error characterisation of the first one needs to be based on existing literature since it concerns several data providers of different processing levels (ground segments, space agencies, etc.) as well as different hardware and software (Mittaz et al., 2019).

Uncertainty characterisation has been increasingly demanded for the last decades by different strata of the EO data user community and, hence, several attempts were done to provide an uninterrupted chain of error propagation at every processing level. For example, in the case of the Moderate Resolution Imaging Spectroradiometer (MODIS) Level 1B product, the MODIS uncertainty index was developed by NASA to provide an estimation of the uncertainties related to that level https://mcst.gsfc.nasa.gov/sites/default/files/file_attachments/M1054E_PUG_2017_0901_V6.2.2_Terra_V6.2.1_Aqua.pdf, last accessed July 2021). For Sentinel data, Gorroño et al. (2017) designed a tool named Sentinel-2 Radiometric Uncertainty Tool (S2-RUT) that aims to estimate the radiometric uncertainty associated with each pixel in the Top-of-Atmosphere (TOA) reflectance factor images provided by ESA. Within the CCI programme, the responsibles of one of the ECV variables, the Sea Surface Temperature (SST) data producers are making important progress in the uncertainty characterisation of the whole retrieval process by propagating the errors associated to Level-1 data of AVHRR to the final SST product (Merchant et al., 2017; Mittaz et al., 2019). Besides, different approaches (prognostic and diagnostic) for the characterisation of uncertainty in aerosol retrievals were recently reviewed by Sayer et al. (2019).

The antecedents of uncertainty quantification for BA products are limited. Giglio et al. (2010), for example, developed the BA uncertainty quantification used on the BA Global Fire Emissions Database version 3 (GFED3) and later on the version 4 (GFED4) products. The GFED products provide monthly aggregate BA extents at 0.25-0.5° spatial resolution based on the MCD64 BA product, which gives the date of detection at 0.5 km spatial resolution derived from MODIS imagery coupled with active fire observations of the same sensor (Giglio et al., 2009; Giglio et al., 2018a). The uncertainty of GFED products is expressed as a standard error of the BA extent estimated in the grid cell, and it is modelled with a linear regression of the burned patch residuals versus the actual extend of burned patches. In their study, the authors computed the per-patch residuals using reference data produced manually from Landsat imagery, at sample sites located in Siberia, Africa and North America.

Within the Fire_cci project, three main analyses should be cited as precursors of the current document. Padilla and Chuvieco (2014) studied different general approaches that are used in error characterisation, although not all of them were found to be suitable for BA. They stated that the analytical based approach described in GUM 2008 could not be used in the Fire_cci project given the dependence of the products on complex spatiotemporal functions and decision trees. This is applicable to current Fire_cci algorithms since they detect BA using machine learning or spatiotemporal thresholding approaches, for which error propagation is considered unfeasible (Merchant et al., 2017). Similarly, the use of Monte Carlo simulations (Crosetto et al., 2001; Crosetto and Tarantola, 2001; GUM-101, 2008) to characterise the BA uncertainty was rejected by the authors. They considered that the input error simulations are frequently very complex, as they must emulate the autocorrelations between errors, which may vary in time and space. Furthermore, they stated that the Monte Carlo approach needs very large computational resources and the

knowledge of the probabilistic distributions of the input data errors, which are not available in the Fire_cci data. Therefore, the authors used an inductive approach based on validation data and regression analysis, which was commonly used in land cover maps to analyse the probability of misclassification (Burnicki, 2011; Smith et al., 2003; van Oort et al., 2011).

The problem with this approach is that it was designed to provide a probability of burn only for the burned pixels and not, as it is required by definition and explicitly by the end-users, for both burned and the unburned classes, since every observed pixel has an associated uncertainty. A natural concern that also arises from these approaches is the quality of the sampling provided by such validation datasets, since even the larger and more systematic validation efforts may still provide only a limited sampling of the true uncertainties (Brennan et al., 2019).

Lewis et al., (2018) performed an assessment of the uncertainty estimates generated by several BA algorithms within the Fire_cci project, starting from the assumption that a per-pixel estimate giving the probability that a pixel should be labelled 'burned' was provided. The actual aim of that assessment was to give some advice to algorithm developers about how to improve their uncertainty estimations, also to show some examples on how the uncertainty was propagated through different processes such as BRDF or ratio of burn indices. Again, those propagation methods were more in line with the "general law of error propagation" described in GUM 2008, which, as was previously stated, it is not feasible to apply in the algorithms that have been and are being developed within the project. Conversely, the same document presented an approach to aggregate the uncertainty estimates from the pixel product to the climate modeller's 0.25-degree grid scale, which was acknowledged as valid in most cases of the current project (Section 5).

Finally, Brennan et al. (2019) estimated theoretical uncertainties for three widely used global satellite-derived BA products using a multiplicative triple collocation (TC) error model that was first described by Stoffelen (1998). This approach considers three observational records X_1, X_2, and X_3 of a variable with an unknown but true value T. The TC error model specifies that each observational record may be related to the truth via a linear measurement equation. The system can be solved based on three initial assumptions, i.e. each product has zero mean residual errors and the errors are uncorrelated with each other and with T. However, a requirement to apply this approach is that the three observation datasets explicitly cover the same temporal and spatial domain. Therefore, the three products were aggregated from the original pixel resolution products to a common sinusoidal grid with a spatial resolution of 1° at the Equator in 16-days periods, which may imply new sources of uncertainty. Obviously, this approach cannot be used in the Fire_cci project since it needs three different BA products to obtain what is called the 'true' value.

Taking all this into account, three main objectives or steps were established for the current deliverable:

1. Provide a detailed overview of the different concepts and common frameworks of uncertainty characterisation and propagation.

2. Describe the different sources of errors encountered in the pre-processing phase as well as their uncertainty characterisation and propagation. Those processes are required to obtain the input data for each BA algorithm developed within the Fire_cci project. In that sense, we will follow the guidelines of Mittaz et al. (2019), which show how to apply the principles of metrology to deal with those issues in EO.

3. Design an uncertainty characterisation and propagation framework for the BA algorithms. In the case of threshold-based algorithms the idea, whenever possible, is to use validation data along with Monte Carlo simulations based on the knowledge of the error's correlation and distribution obtained in the previous step 2. For those algorithms that use machine learning approaches a simple way of characterizing uncertainty is to use the classification probability, which describes the likelihood that a pixel belongs to the burned class.

4. Provide an aggregation methodology to propagate the uncertainty from the pixel to the grid product depending on the specifications of each algorithm.

As a result, and similarly to the previous Fire_cci projects, an important practical output of the error characterisation called uncertainty layer has been generated for the pixel and grid product. For the pixel product, the uncertainty is expressed in probabilistic terms, as the probability that a pixel is really burned. This layer was named confidence level in a previous version of the Fire_cci products (Chuvieco et al., 2017). For the grid product, uncertainty is expressed as a standard error of the total burned area for each grid cell, as requested by the Climate Research Group (CRG) (Chuvieco et al., 2017).

2.2 Terminology

Several terminological concepts are extremely important to consider before starting uncertainty analysis. In this section, we are describing the different terms in relation to the process based on the GUM (2008) and Mittaz et al. (2019):

- **Measurand:** a particular quantity subject to measurement. All ECVs involve an indirect measurement process by applying several transformations to raw signals.
- **Accuracy:** refers to the closeness of agreement between a measured value and the true value of a measurand.
- **Precision:** denotes the degree to which further measurements or calculations show the same or similar result. A measurement with a small random uncertainty is said to have high precision.
- **Calibration:** defined as the process of converting the raw signal recorded by the satellite to the measurand. Examples include converting raw counts to a radiance or brightness temperature. The calibration process is normally defined by an algorithm and a set of calibration coefficients.
- **Discrepancy:** the difference between the measurement and the validation value. A small average discrepancy with respect to the root-sum-square of the measurement and validation value uncertainties is indicative of accurate measurement, but it could also result from a fortuitous cancellation of error terms.
- **Bias:** an offset (additive) or scaling factor (multiplicative) that affects all measurements made by a particular instrument. The bias may be estimated, in which case it can be corrected for (a correction) or may be an unknown error. In practice, it refers to the mean value of the discrepancy.
- **Correction:** denotes any operation of adjustment made to correct for a known bias. This may have a functional form (e.g. a straight line) with multiple correction parameters (e.g. an offset and slope). Note that even after correction there will always be a residual, unknown error.
- **Error (of measurement):** the process of measurement is never exact, and the difference between a measured value and the 'true value' of the measurand is called the error.

- **Random (aleatoric) errors:** are errors manifesting independence: the error in one instance is in no way predictable from knowledge of the error in another instance. A complication arises in EO imagery when one instance of a parameter in the radiance measurement function is used in the calculation of the Earth radiance across many pixels. That component of the error in the radiance image is then correlated across pixels, even though the originating effect is random.

- **Structured errors:** arise from effects that influence more than one measured value in the image but are not in common across the whole image. The originating effect may be random or systematic (but acting on a subset or locality of pixels), but in either case, the resulting errors are not independent and may even be perfectly correlated across the affected pixels. Since the sensitivity of different pixels or channels to the originating effect may differ, even if there is a perfect error correlation, the error (and associated uncertainty) in the measured value can differ in magnitude. Structured errors are therefore complex, and at the same time, important to understand because their error correlation properties affect how uncertainty propagates to higher-level data. A structured random effect would refer to an effect that is unpredictable in terms of origin while leading to a predictable pattern of correlated errors across measured values in an image.

- **Common errors:** These errors are constant (or nearly so) across the satellite image and may be shared across the measured radiances for a significant proportion of a satellite mission. Common errors might typically be referred to as biases in the measured radiances.

- **Systematic errors:** arise from unknown effects that could in principle be estimated rather than from chance processes, and that influence many measured values, including, but are not limited to, effects that give rise to constant error for a significant proportion of a satellite mission – i.e., biases, for which the structure is a simple error in common. Generally, this type of errors can be corrected for if we had sufficient information to do so. However, our knowledge of the systematic error is imperfect, and so our correction, although the best we can do with the information available, will be imperfect. In terms of correlation properties across an image, therefore, effects that are systematic in origin give rise to either structured or common errors. Systematic errors, therefore, "average out" slowly or not at all across many measured values; systematic effects may be operating at the same time as other types of effect, in which case only a component of the total error is systematic; an example of a systematic effect is a mischaracterised calibration target.

- **Uncertainty (of measurement):** a parameter, associated with the result of a measurement that characterizes the dispersion of the values that could reasonably be attributed to the measurand. The estimation of uncertainty is given in the same unit as the measured value. This is generally written as the standard uncertainty $u(x)$ where x denotes the measured value.

- **Standard uncertainty:** it describes the standard deviation of the probability distribution describing the spread of possible values.

- **Consistency (of CDRs):** can be defined as the compatibility of the relevant characteristics of two or more climate data records (CDRs) (e.g., patterns, variability, and trends) with a reference (represented by a physical equation, a model, or a fiducial reference) within their combined uncertainties.

2.3 Uncertainty analysis

The ultimate goal of uncertainty propagation is to evaluate the reliability of the model outputs, to represent the model output variability in a compact way. Let's consider the modelling function of the measurand Y from the vector of input quantities $X = \{X_1, X_2, \dots X_n\}$:

$$Y = f(X) + \Delta \tag{1}$$

The term Δ is an input quantity introduced to represent any inadequacy of the function f to represent all phenomena that affect the measurand. In the GUM, the convention is for estimates to be represented with the lower-case characters corresponding to the quantities written in upper case. Thus the estimated value of Y will be modelled hereafter as:

$$y = f(x_1, x_2, \dots x_n) + \delta \tag{2}$$

The evaluation of the uncertainty in y can be propagated through the measurement model of these uncertainties (or, strictly, distributions). In practice, the process of uncertainty analysis involves the determination of uncertainty budget, which is the list of random and systematic errors components that contribute to the uncertainty in measurement results.

In the best-case scenario, where the relationship between input variables and the measurand can be algebraically written, the GUM recommends propagating the uncertainty using an analytical method well known as the "Law of Propagation of Uncertainty" (LPU). The LPU propagates standard uncertainties for the input quantities through a locally-linear first-order Taylor series expansion of the measurement function to obtain the standard uncertainty associated with the estimate y of the measurand (higher order approximations can be applied if necessary). Then, the uncertainty of the measurand can be written in matrix notation as:

$$u^2(y) = C \, S_x \, C^T \tag{3}$$

where S_x is the covariance matrix of input quantities and C is the vector of sensitivity coefficients, which provide a measure of how sensitive the measurand, Y, is to a change in a given input quantity, X_j. They propagate the uncertainty associated with an input quantity into an uncertainty associated with the measurand (and in the units of the measurand). The simplicity of the analytical expression of the measurand allows to get the sensitivity coefficients based on the partial derivative of the measurement function with respect to the term that this uncertainty applies to, evaluated at the estimates of the input quantities. Thus, the sensitivity coefficient of a given input quantity, X_j can be defined as:

$$C_j = \frac{df}{dx_j} \tag{4}$$

The concept of correlation error is very important to account for in that process. For this reason, the total uncertainty of the measurand is defined by the covariance matrix of the different input variables denoted S_x:

$$S_x = \begin{bmatrix} u^2(x_1) & u(x_1)u(x_2) & \dots & u(x_1)u(x_n) \\ u(x_2)u(x_1) & u^2(x_2) & \cdots & u(x_2)u(x_n) \\ \vdots & \vdots & \ddots & \vdots \\ u(x_n)u(x_1) & u(x_n)u(x_2) & \cdots & u^2(x_n) \end{bmatrix} \quad\quad (5)$$

where $u(x) = \{u(x_1), u(x_2), \dots u(x_n)\}$ is the vector of input uncertainties.

Errors and their uncertainty budget can be summarised in a consistent manner using the uncertainty analysis diagram. Figure 1 illustrates an example with three input variables. The errors affecting these variables are described by their probability distribution functions (PDF) that allow deriving the uncertainty as a standard deviation. The term 0 in the diagram denotes the best estimate of δ, which is the expectation of Δ.
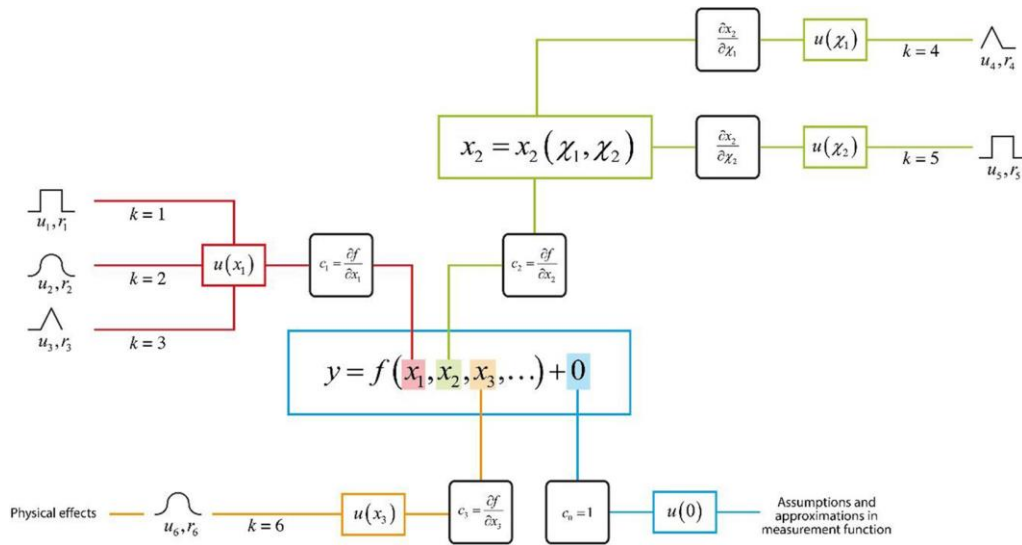
Figure 1. Uncertainty analysis diagram. Source: Mittaz et al. (2019)

In the case of highly non-linear models with more complex probability distributions, and models that cannot be written algebraically, which are very common in geophysical variables retrieved from EO data, the GUM recommends using Monte Carlo simulations to shape the output probability distribution for the measurand. In contrast to the analytical LPU method, Monte Carlo may be regarded as a numerical approach that uses algorithmically generated pseudo-random input values. Then, the values are forced to follow some prescribed probability distributions of input variables in order to get a finite set of random draws from those distributions. Monte Carlo propagates the sets of input values afterwards through the measurement function to obtain a set of output values regarded as random draws from the probability distribution of the measurand. This process is repeated a sufficiently large number of times to produce a set of simulated results as output. In the end, the mean and standard deviation of these output results are considered as the best estimates of the measurand and its standard uncertainty (Farrance and Frenkel, 2014).

Monte Carlo method evaluates the measurement model in each trial; this means that it is very costly in terms of computational load. This issue is very critical in the case of very large EO datasets, which is the reason why it is often used in combination with LPU. For instance, it can be applied to determine the uncertainty for a particular quantity, which is then used as an input to LPU in a subsequent uncertainty analysis.

2.4   Status of uncertainty analysis at the level of CCI programme

The question of how to derive and present uncertainty information in CDRs has received sustained attention within CCI programme (Hollmann et al., 2013). For this reason, all ECVs deliverables within the frame of the programme were requested to include a detailed description of the sources of uncertainty in the different processing stages, and a consistent manner of uncertainty propagation. The documentation of most of ECVs covers these aspects in their error characterisation reports (except in some cases, e.g. Snow CCI). However, the practical application of uncertainty propagation methods from level-0 until level-3 and level-4 products is still challenging. For instance, a detailed description of sources of errors encountered while processing SAR and optical data to deliver the Land Cover CCI product is given, along with the approaches of uncertainty estimation. However, only uncertainty derived from classification algorithms is delivered with pixel product. Moreover, in the case of the grid product, the results weren't good enough for climate modelling. It was thus proposed to express the uncertainty in the form of a new set of "alternative" global land cover maps. These "alternative" maps should give, on a per-pixel basis another possible class when relevant and using them in the models instead of the original LC maps (ESA, 2015). Concerning Lakes CCI, the uncertainty estimation of lakes water level parameter (LWL) at a given time is derived from the standard deviation of the individual lake water height along the track, but uncertainty propagation is not reported. The uncertainty estimates of other delivered variables in that product are mainly carried out using reference data (Vickers et al., 2020). On the other hand, uncertainty propagation was possible in the case of Biomass CCI by applying the LPU to the measurement model and by approximating uncertainty of S1 backscatter as a function of the equivalent number of looks (Santoro et al., 2019). In general, Researchers commonly encounter datasets where uncertainty information is generic, misleading or absent (Merchant et al., 2017), which makes the correct uncertainty propagation of these ECVs a challenging task, and sometimes impossible.

For the Fire_cci project, we have undertaken a similar approach, by considering the main factors of uncertainty associated to the different phases of the processing chain: firstly, focus on the pre-processing (geometric and radiometric correction: section 3) and secondly, on the burned area detection (section 4). We have developed different procedures adapted to the sensors used as input and the algorithms for the different Fire_cci BA products.

3   Uncertainty characterisation through pre-processing steps

3.1   General overview

In a laboratory, the uncertainty is characterised by the statistical evaluation of repeated measurements. In EO, this is not possible due to the variation in sensor state, viewing geometry, and natural geophysical variability. Besides, the atmosphere modifies the TOA radiance observable from space by processes of scattering, absorption, and emission. The impact of the atmosphere on the radiance depends on the vertical profile of radiatively active gases, aerosols and clouds. Surface changes caused by variations in moisture, temperature, or vegetation phenology, among other factors. These variations also affect reflected radiance and may confuse the detection of actual cover changes, such as those caused by fire disturbances.

To understand the uncertainty propagation throughout the process of measurement and posterior transformations of the obtained data a key term is the level of processing of an

EO product. Processing levels reflect both distinct computational stages in handling data streams downlinked from satellites and the different institutional arrangements for creating products at different levels (Mittaz et al., 2019). Although it is not yet standardized among agencies and communities that provide those products, typically five levels can be differentiated:

- Level-0 (L0): This level includes the raw telemetry that is downlinked by a ground receiving station, which comprises a mix of scientific observations together with engineering data. Transforming L0 data to scientifically useful products is complex. One of the main sources of uncertainty is the digitisation carried out to transform the analogic signal into digital. The raw sensor data is binary (10 to 16 bits), but such digitisation is coarse compared to laboratory metrology. That binary representation of the raw sensor values places a fundamental lower limit on the uncertainty present in the calibrated radiances. For example, 10-bit digitisation corresponds to 0.1% resolution of the range (Mittaz et al., 2019). In addition to this source, L0 processing involves estimating the satellite orbit and the origin of the measured radiances projected onto Earth's surface (known as geo-locating). In this case, the uncertainty increases depending on the quality of the orbit information, i.e. if it is Near Real Time (NRT) or Non-Time-Critical (NTC) improved geo-location.

- Level-1 (L1): It includes calibration parameters to map the counts into radiance and the derived calibrated radiances. Auxiliary data locate the radiances in time, latitude, and longitude and provide information related to satellite and solar zenith and azimuth angles. At this level, the main sources of error are the calibration parameters (gain and offset) that are part of the measurement function. In this case, this function allows the previously mentioned mapping of counts into radiance. Changes in those parameters are to be expected as the sensor's space environment changes and the sensor degrades. This means that the uncertainties associated with measured radiances will evolve. Typically, after 3-10 years a sensor will fail or will be decommissioned. In the case of multi-decadal datasets, the sensors are supposed to have an identical spectral response but they have significant differences in their Spectral Response Function (SRF).

- Level-2 (L2): This level normally involves the inverse estimation from radiances of one or more geophysical variables. The retrieval algorithms used in EO for that purpose are highly varied. As in previous levels, no input uncertainty information is included at this level. Users of L2 often interrogate pixel-level quality indicators for indications, but quality indicators change from one product to another. In that sense, in ESA CCI they reached a consensus with regards to the best practice for geophysical products (Merchant et al., 2017).

- Level-3 (L3): They are gridded products, made by aggregating L2 values in space and/or time on a regular space-time grid. Therefore, it implies averaging the L2 (sometimes weighted). Assuming the sampling or the aggregation is independent of the variability, the resulting sampling error has an expectation of zero. But the error distribution has a width that is not negligible compared to the uncertainty of the value.

- Level-4 (L4): It typically includes complete gridded gap-free in space and time information, which in many cases involves interpolation, and hence, a specific error propagation.

Officially, no complete traceable analysis and propagation of uncertainty from L0 to L4 exists for any current EO processing chain (Mittaz et al., 2019). This means that algorithm developers have no chance to include as input the error characterisation of previous levels and, therefore, to generate a proper error characterisation of their own processes. The following sub-sections try to characterise from existing literature the different error sources

of those previous levels to somehow show the diversity and nature of the errors that should be taken into account.

3.2   Terra-MODIS Level 2G

Three products are used as main inputs for the FireCCI51 algorithm (Lizundia-Loiola et al., 2020) related to the Terra satellite's MODIS sensor: the MOD09GQ product, which provides an estimate of Red and Near-Infrared (NIR) surface reflectance at 250 m spatial resolution, the MOD09GA product, which provides, along with the rest of the reflectance and emissive bands, the state QA flags used by the algorithm at 1000 m spatial resolution, and the MCD14ML active fire product. This latter product includes also Aqua satellite data and will be covered in the ancillary data section (Section 3.7), as it is only used for guiding the training phase.

Four MODIS processing levels are differentiated based on the MOD09 user guide (Vermote et al., 2015). L0 data is raw satellite data that feeds L1 data that has been radiometrically calibrated, but not otherwise altered. L2 data is L1 data that has been atmospherically corrected to yield a surface reflectance product. L3 data is L2 data that has been gridded into a map projection, and usually has also been temporally composited or averaged. L4 data are products that have been put through additional processing. All data up to and including L2 are in an ungridded orbital swath format, with each swath typically cut into small segments, or granules, to facilitate processing. Data at L3 and up are geo-located into a specific map projection, with the geo-located products typically in a set of non-overlapping tiles. The L2G-lite format, consisting of gridded L2 data, was developed as a means of separating geo-locating from compositing and averaging.

Several scientific teams are responsible for the quality of the MODIS products. The MODIS Land Quality-Assessment team, for example, evaluates and documents the scientific quality of the MODLAND products with respect to their intended performance. The MODIS Characterization and Support Team (MCST) is responsible for developing and maintaining the MODIS calibration product (L1B algorithm), which is a precursor to every geophysical science product. In this last case, there has been an attempt to characterise the uncertainty that affects the L1B creation process through the inclusion of uncertainty data on a pixel-by-pixel basis. The reflective solar bands uncertainty algorithm is based on characterization of the Earth View (EV) scene reflectance (Xiong et al., 2013). To compute it, the uncertainty of the Solar Diffuser/Solar Diffuser Stability Monitor calibration, the lunar calibration, EV response trending, the temperature correction and the scene dependent instrument noise are taken into account.

In the following step, the L1B is corrected for the effects of atmospheric gases and aerosols to obtain the L2 product (MOD09). Thus, an estimate of the surface spectral reflectance for each band as it would be measured at top of the canopy if there were no atmospheric scattering or absorption is obtained. Band quality control information for the correction is also generated (e.g., flags denoting if ancillary data is unavailable, if L1B data is faulty, etc.). However, this layer does not represent actual uncertainty propagation data, but a general band quality of the Atmospheric Correction (AC).

Finally, daily MOD09 L2 data is aggregated in standard MODIS sinusoidal tiles (MOD09GA and MOD09GQ L2G-lite). The number of daily observations at each pixel is determined not only by the number of orbits at that location (one at the equator and up to 15 at the poles), but also by the spread of observational coverage of off-nadir pixels. After identifying all the observational values for each location, the most suitable reflectance value is selected for each pixel based on observational coverage and view angle, and

whether the observation is flagged as cloudy, clear, containing high aerosol or low aerosol, or cloud shadow.

The L2G-lite products, which are the inputs for FireCCI51, include a layer of the band quality that cannot be considered a proper uncertainty layer. Although the original uncertainty of the L1B product is somehow taken into account when setting the quality of a pixel, it provides qualitative information instead of a quantitative description of how errors are propagated and distributed. No information is officially provided in the products regarding the error propagation through the AC, projection of the data or how the ranking that is used to select the best observation affects uncertainty.

In that sense, Vermote et al (2008) made an estimation of the sensitivity of surface reflectance to uncertainties in input key atmospheric parameters using a theoretical error budget. Such a budget was created based on the simulation of a number of atmospheric and geometrical scenarios (Vermote and Saleous, 2006). In that study, uncertainties from the instrument calibration ($\pm2\%$), atmospheric pressure (±10 mb), water vapour content (±0.2 g/cm^2), ozone content (±0.02 cm·atm), retrieved Atmospheric Optical Thickness (AOT) values (resulted from the aerosol inversion), and selection of the aerosol model (urban polluted, smoke low absorption, smoke high absorption, or urban clean) were considered. They found that the overall accuracy of surface reflectance varies depending on the band and AOT. Under clear atmospheric conditions, they stated that it does not exceed 0.006 in reflectance unit. The MODIS product theoretical uncertainty bars were set to $0.005+0.05\rho$ for the surface reflectance under favourable conditions (no high aerosol). To check the quality of the MOD09 they analysed a year of Terra data (2003) collected over 150 AERONET sites. The results showed that the average percentage of observations that lay within the theoretical uncertainty for bands 1 (645 nm), 2 (870nm), 3 (470 nm), 4 (550 nm), 5 (1240 nm), 6 (1650 nm), and 7 (2130 nm) was equal to 88.66%, 94.34%, 50.52%, 79.34%, 96.50%, 97.87%, and 98.62%, respectively. The band 2 NIR, which is used by the FireCCI51, showed one of the highest proportions of "good" observations.

Regarding the state information that is used from the MOD09GA, although it is unclear the origin of all bit values included on it, it seems that the main source is the MOD35 developed by the MODIS Cloud Mask Team (Ackerman et al., 2010). This product, which is based on L1B radiance data, assumes that its inputs are calibrated and quality controlled and no propagation of its uncertainties is made. In the process of identifying pixels, several thresholds are used and hence, as one approaches the threshold limits, the certainty or confidence in the labelling becomes more and more uncertain. For that reason, a confidence layer is provided that tries to be an alternative to uncertainty considering it as a function of how close the observation is to the thresholds.

3.3   Sentinel 3-Synergy Level 2

This product is the input for the FireCCIS310 algorithm, which will be available for the years 2019 and 2020. One unique product is used as main input related to the Sentinel-3 (S3) satellite's OLCI and SLSTR sensors: the SY_2_SYN product, which provides an estimate of Visible, NIR and Short Wave Infrared (SWIR) surface reflectance at 300 m resolution, along with the rest of the information, e.g. quality flags or geometry.

Several processing levels, similar to those established by MODIS, can be differentiated based on the S3 OLCI, S3 SLSTR and S3 Synergy (SYN) Product User Guides (PUG, https://sentinel.esa.int/web/sentinel/user-guides, last accessed July 2021). L0 data is raw satellite data that includes time sorted and annotated data from Instrument Source Packet (ISP). These data, the orbit scenario file and several auxiliary data files feed the L1 EO

processing chain. The derived product is called L1B and provides radiometrically calibrated, geo-referenced and annotated radiances. These first two levels are common but independent for both OLCI and SLSTR sensors and, hence, two L1B products are obtained, respectively. In the following step, these two products are used to feed and generate the internal SYN L1 product (SY_1_MISR). This process aims to project all OLCI and SLSTR bands on the same SYN reference grid (i.e. the OLCI acquisition grid), using the inter-instrumental misregistration estimated for both sensors' bands. Finally, in L2 that internal information is atmospherically corrected to generate surface directional reflectances stored in the SY_2_SYN product. As in the case of MODIS, all data up to and including L2 are in an ungridded orbital swath format, with each swath typically cut into small segments, or granules, to facilitate processing. However, in the case of SYN there is not any processing level equivalent to the L2G-lite of MODIS. Therefore, SYN L2 information is aggregated into daily non-overlapping tiles of 10x10 degrees by Brockmann Consult (BC) for internal use.

The quality of the S3 products is the responsibility of the full Ground Segment and mainly of the Payload Data Ground Segment (PDGS). It operationally generates the user products and distributes L0 raw products, processed L1 products and derived L2 products. However, there is no official uncertainty propagation process integrated through the different levels.

In the case of the OLCI L1B products only quality flag information (saturated radiances, dubious locations, where cosmetic was applied, bright pixels, sun glint risk, etc.) is provided within the product (ACRI-ST IPF Team, 2017), along with some information related to the observation environment, such as the geometry tie points and meteorological tie points. It is supposed that an error estimate band of the radiance is included, but no such layer was found in the downloaded data from the open hub (https://scihub.copernicus.eu/dhus/#/home, last accessed July 2021). A similar approach is followed for SLSTR, although in this case the previous information is extended due to the availability of SWIR and Thermal Infrared bands (ACRI-ST IPF Team, 2018). Each SLSTR band is provided along with a quality data layer that contains estimates of the random signal noise per scan line derived from the on-board calibration sources (the VISCAL and the blackbodies), and correlated radiometric uncertainties as a function of scene radiance or/and brightness temperature (BT) derived from the pre-launch calibration.

Although there is no proper uncertainty propagation information within the official products, some attempts to assess the per-pixel uncertainty of both sensors L1B products have been found in the literature.

On the one hand, Hunt and Nieke (2016) developed a software tool to determine the per-pixel uncertainty of the OLCI L1B images. It was developed following the approach outlined in GUM (2008) and was created with the goal of integrating it in the S3 Toolbox (now called Sentinels Application Platform, SNAP) software package as a plugin. Similarly to Mittaz et al. 2019, they first describe the OLCI measurement model of the L1B radiances to characterise the contributors of the overall uncertainty model. Then, they defined an uncertainty model as a function of the total signal (noise), radiometric calibration (diffuser BRDF estimated at 0.3%, diffuser alignment estimated at 0.31%, diffuser ageing, calibration diffuser stray light estimated at 0.08%, calibration camera stray light estimated at 0.2% and calibration speckle estimate at typically 0.1%), non-linearity (CCD and ADC non-linearity), dark signal (offset compensation and dark stability error), smear (smear gain contributions) and stray light contribution. Hunt and Nieke (2016) showed uncertainties around 0.76% for a specific preliminary example for a region in Northern Sahara.

On the other hand, Etxaluze and Smith (2019) have recently developed a tool called MapnoiS3 to allow users of SLSTR L1B data to derive per pixel uncertainty estimates for both radiance and thermal channels using the information contained in the L1B product and additional auxiliary data files (ADF). Using the random signal noise and correlated radiometric uncertainties included in the quality layers they are able to generate a new NetCDF file where the scene radiance/BT uncertainty, noise equivalent radiance/BT and, in the case of BT, the partial derivative of the radiance as a function of temperature are included. The contributors to the total uncertainty slightly differ from radiance to BT. In the first case, the total radiometric noise is considered to be composed of the noise related to the light intensity level (shot noise) and the electronic noise (dark current, amplifier noise, reset noise, digitisation). However, Etxaluze and Smith (2019) stated that at the high-level quantisation of SLSTR the digitisation, amplifier and reset noise are insignificant. In the latter case, the radiometric noise is considered to be composed of the noise equivalent differential temperature per scan line.

In the following processing steps, regarding the SYN product, two consecutive products are generated: the L1C and the L2. The former is only for internal use and it is not resampled to a specific surface grid or projection, but includes all the necessary misregistration information so that any user-defined projection or gridding can be performed at a higher-level (S3 Team, 2011). There is not any mention of the uncertainty in the whole L1C Algorithm Theoretical Basis Document (ATBD, https://sentinel.esa.int/documents/247904/349589/SYN_L2-3_ATBD.pdf/8dfd9043-5881-4b38-aae5-86fb9034a94d?t=1371548198000, last accessed on August 2021). The latter contains atmospherically corrected surface directional reflectance referenced at OLCI geometry that is generated based on the L1C product. However, in this case, there is a dedicated section to the retrieval error estimate of the L2 product (North and Heckel, 2010). Two main contributors are analysed there: the error in Aerosol Optical Depth (AOD) and the surface reflectance error as a function of the previous error, sensor noise and estimated error in the radiative transfer model. This last error is computed for each spectral band and included as a separate layer along with the reflectance layer. It is offered in the same relative units of reflectance (%). Since the ATBD was published previous to the development of the above mentioned L1B uncertainty tools, there is not any mention of them. The lack of official error distribution characterisation at previous levels along with the lack of information of that process questioned the reliability of those layers. However, we consider the inclusion of a per-pixel uncertainty layer a great improvement and its use will be studied.

In the last step, the SY_2_SYN L2 product is aggregated temporally into daily images and spatially into 10x10 degree tiles. The scientific content of the product is not in any case altered, but the best observation is selected when more than one valid observation is available for the same pixel. Similarly to what is done with the MODIS L2G lite product, the best observation is considered when OLCI zenith angle (OZA) is minimum. The values of all the layers of the selected image are kept in the final aggregated daily product including the uncertainty associated to that pixel.

3.4   NOAA-AVHRR LTDR Level 2G

FireCCILT10 (Otón et al., 2019) uses the AVH09 Surface Reflectance Product (Version 5) that is provided by the Land Long Term Data Record (LTDR, Pedelty et al., 2007) project. It is based on the Global Area Coverage (GAC, 4x4 km) data with a final spatial resolution of 0.05º degree (≈5 km). AVH09 covers four decades (1981–present) and it includes up to two sensors (AVHRR2 and 3) on board seven different NOAA satellites (7, 9, 11, 14, 16,

18, 19). It provides Red and NIR Surface Reflectance, and TOA BT, all of them used by the FireCCILT10 algorithm.

First, L0 data is converted on AVHRR L1B (Robel et al., 2014), which contains the main geophysical parameter reflectivity (for channels 1, 2 and 3a) and calibrated radiances (for channels 3, 4 and 5) (EUMESAT, 2011). At this point, the uncertainty is affected by several features, such as the Platinum Resistance Thermometer (PRT) noise and bias, earthshine, Temperature gradient, PRT representation, detector noise, digitisation, amplifier, SRF, time mismatch, space mismatch, azimuthal asymmetry, solar contamination, pre-flight characterisation, degradation and instrument temperature dependence. Some of these sources directly affect thermal gradients across the internal calibration target and the estimation of calibration parameters. Regarding the solar reflective bands, the calibration is different for the visible and the infrared (IR) channels (EUMESAT, 2011). Although Red and NIR bands are calibrated prior to launch (Robel et al., 2014), the calibration is applied to the data afterwards with linear decay because there is not effective on-board calibration (Holben et al., 1990). The uncertainty in the calibration is estimated to be of the order of 5% (Robel et al., 2014). In the case of the emissive bands, BT is calibrated in flight with an uncertainty estimated at ±0.1 K (Trischenko et al., 2002). In both cases, several problems were observed due to the combination of different sensors and satellites. In general, AVHRR3 was seemed to have better radiometer consistency than AVHRR2 (Trischenko et al., 2002). Besides, it was observed that the calibration of the different satellites generated a range of variability in the measurements that, in the case of BT, could be more than 0.5 K (Trischenko et al., 2002; Mittaz et al., 2019). A quality decrease and systematic degradation of the radiometric sensitivity as a function of time and launch processes was observed as well. This degradation was quantified in ≈1-3% per year by Uprety et al. (2011) in the case of the solar reflective bands and below 1% in the case of BT. These generate a reduction in radiometric resolution over time and, in the case of BT, expands the upper limit of the measured BT. In any case, a proper characterisation of the uncertainty generated by all those sources cannot be done since there is no access to the pre-launch measurements and to the original manufacturer's data propagation (Mittaz et al., 2019).

In the following step, L1B is transformed in GAC, reducing in real time the spatial resolution of the acquired image (EUMESAT, 2011). Four out of every five samples along the scan line are used to compute one average value, and the data from only every third scan line are processed. As a result, the spatial resolution of GAC data near the subpoint is 1.1 km by 4.4 km with a 2.2 km gap between pixels across the scan line (Robel et al., 2014; EUMESAT, 2011). Bulgin et al. (2016) estimated the uncertainty of this averaging process on about 0.04 K.

Finally, GAC is processed to obtain AVH09 L2G product, after applying some improvements. These improvements include: radiometric in-flight vicarious calibration for the visible and near infrared channels, inverse navigation to relate an Earth location to each sensor IFOV, atmospheric corrections for Rayleigh scattering, ozone, and water vapour, aerosol correction, and BRDF corrections used in MODIS processing (El Saleous et al., 2000; Pedelty et al., 2007). Calibration is a critical issue for applications using multiple sensors, including multi-decadal data analysis (Mittaz et al., 2019), and AVH09 has demonstrated 1% calibration accuracy for the visible/NIR bands (Pedelty et al., 2007). However, an uncertainty characterisation and propagation through the above-mentioned processing levels is still needed.

3.5 Sentinel 2-MSI Level 2

The Sentinel-2 (S2) MSI L2 processing chain generates, using as input the TOA L1C orthoimagery, Bottom-Of-Atmosphere (BOA) corrected reflectance orthoimages. Additionally, an AOT map, a Water Vapour (WV) map and a Scene Classification Map (SCM) together with Quality Indicators (QI) for cloud and snow probabilities are generated. S2 bands at 10, 20 and 60 m spatial resolution are required for L2 processing. Spectral bands 2, 3, 4, 8 (Table 1) as well as a True Colour Image (TCI) and AOT and WV maps are provided at 10 m. Spectral bands 2-7, 8A, 11, and 12, and resampled TCI, AOT and WV are provided at 20 m. Band B8 is omitted as B8A provides more precise spectral information. All components of the 20 m product are resampled to 60 m as well. The cirrus band 10 is omitted, as it does not contain surface information.

Table 1. Central wavelength (nm) of each band of the MSI aboard Sentinel-2 A and B.

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

The Sen2Cor (Main-Knorn et al., 2017) processor is a combination of state-of-the-art techniques for performing atmospheric as well as topographic corrections which have been tailored to the S2 environment together with a scene classification module. The scene classification algorithm allows detection of clouds, snow and cloud shadows and generates a map that contains three different classes for clouds (including cirrus) together with six additional classes: shadows, cloud shadows, vegetated, not vegetated, water and snow. Such a classification is an important information source for uncertainty as it may be used to mask out the pixels covered by clouds and their shadows and then not classify them. However, such a classification also has intrinsic uncertainty related which should be considered. The Shuttle Radar Topography Mission Digital Elevation Model (SRTM DEM) is also used when pre-processing the S2 data. The SRTM DEM has vertical errors between 4-6 m (Goncalves and Fernandes 2005). Such errors may increase the uncertainty when mapping BA, especially over steep slopes.

Finally, BA mapping uses datasets acquired by different satellites (S2 A and B), which may generate geolocation errors. Thus, it is expected that any geolocation error may be significant at fire borders.

3.6 Sentinel 1 SAR

Sentinel-1 (S1) is a two-satellite constellation (A - since April 2014, B - since April 2016) to provide C-Band SAR data continuity following the end of ERS-2 and Envisat missions. 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 instrument may acquire data in four 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 S1 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). One should notice that nominal temporal frequency is not yet achieved over areas outside Europe and North America and

that areas with frequent seismic activity (e.g., the Andes) are imaged in single polarization mode (VV) for increased spatial resolution. S1 products are released in two L1 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 most commonly available SLC and GRD data are acquired in IW mode. L1 GRD data multi-looked and projected to ground range using an Earth ellipsoid model (typical product size is 1GB for dual-pol IW mode) is considered to map BA.

Speckle appears as a granular interference that inherently exists in coherent image systems as the SARs. The presence of speckle makes SAR imagery very different from optical datasets. As the scatterers are not identical for each cell, the signal fluctuates due to its high sensitivity to small variations in scatterers' relative location and properties. Differences in the magnitude of the signal intensity between two dates may appear due to fire unrelated changes (e.g. changes in the relative position of the scattering elements within a resolution cell due to wind) which influences uncertainty. To meet the accuracy requirements of most SAR-based applications, a large equivalent number of looks (ENL) is often used, with values around 100 being typically advised (Quegan et al., 2001). ENL describes the degree of averaging applied to the SAR measurements during data formation and post-processing (Anfinsen et al., 2009). The usual approach to reduce speckle is filtering in the spatial domain. When multiple intensity images of a scene are available (repeat passes), an attractive way to increase the ENL is by linearly combining the images (multi-temporal filtering). Previous experiments, carried out within Fire_cci Phase 2 Option 3, showed small differences (~4) in ENL between S1 products processed with increasing multi-look factors, i.e., spatial spacing of 20, 30, and 50m. At the same time, Sentinel-1 products showed a significantly larger ENL when combining spatial (multi-looking) and temporal filtering. Spatial aggregation to 40 m (2x10 looks) followed by multi-temporal filtering provided ENL values around 80. Moreover, the SAR instruments aboard of S1-A and B satellites have a radiometric error of 0.38 and 0.36 dB, respectively (Schwerdt et al., 2017). We expect that data uncertainty will have low relevance in the results according to previous studies (Belenguer-Plomer et al., 2018).

Apart from intrinsic sensor-related uncertainty (e.g. absolute calibration), environmental conditions at SAR image acquisition may influence BA detection and mapping. Changes in soil moisture may increase BA detection and mapping uncertainty as it may entirely mask out fire-induced variations of the backscatter coefficient or result in fire unrelated backscatter changes (Belenguer-Plomer et al., 2019). However, such effects cannot be measured with precision since current global soil moisture products have a much coarser spatial resolution than S1 data. Topography also introduces uncertainty when using SAR-based datasets as the SAR backscatter coefficient is modulated by terrain slope thus conditioning the accuracy of BA detection (Kalogirou et al., 2014; Kurum, 2015).

3.7   Ancillary data

Two main ancillary data are used by the different algorithms that are being developed within the Fire_cci project: thermal anomalies and land cover information. Both products are high-level products (L3 and L4, respectively), which means that they are based on the above-mentioned processing levels that do not include proper uncertainty characterisation.

3.7.1 Terra and Aqua MODIS active fires

The first source of thermal anomalies that is used by some of the algorithms developed within the project is the MODIS MCD14ML collection 6 product.

In the case of MODIS, the most basic fire products in which active fires and other thermal anomalies (such as volcanoes) are identified are the MOD14 product based on Terra satellite's MODIS and MYD14 based on Aqua satellite's MODIS (Giglio et al., 2018b). These L2 products are defined in the MODIS orbit geometry covering an area of approximately 2340 × 2030 km in the along-scan and along-track directions, respectively. Regarding the uncertainty characterisation of the active fires, a detection confidence level is provided, which intended to help users gauge the quality of individual fire pixels. This confidence estimates range between 0% and 100% and is used to assign one of the three fire classes (low-confidence fire, nominal-confidence fire, or high-confidence fire) to all fire pixels (Gilgio et al., 2018b). The confidence of each detected fire pixel is calculated as the geometric mean of up to five sub-confidence parameters that are defined in terms of the temperature, the number of adjacent water pixels, the number of adjacent cloud pixels and two standardised variables (Giglio et al., 2016). All the parameters used for the detection are related to the algorithm used to detect thermal anomalies and the uncertainties of the previous processing levels are not taken into account. Besides, no confidence level information is provided for those pixels that were non-detected as fire or filtered in the processing. Thus, this confidence level can only be used as an approximation of the product uncertainty.

The level 3 MCD14ML product is generated based on MOD14 and MYD14. This monthly fire location product contains the geographic location, date, and some additional information for each fire pixel detected by the Terra and Aqua MODIS sensors on a monthly basis. The product is distributed as a plain ASCII (text) file with fixed-width fields delimited with spaces. One of its attributes is the previously mentioned confidence level.

3.7.2 NOAA-20 and SUOMI-NPP VIIRS active fires

Information related to the VIIRS products is significantly lower than for the MODIS ones. VIIRS products are supposed to continue with the MODIS time series so most of the algorithms are being adapted to the new sensor. Therefore, the conclusions reached for MODIS active fires in the previous section (Section 3.7.1) can be considered as the best scenario for VIIRS active fire uncertainty characterisation.

Currently, there are two satellites providing active fire products based on VIIRS: the Suomi-NPP and the NOAA-20 (JPSS-1). Recently, the Fire Information for Resource Management System, FIRMS (https://firms.modaps.eosdis.nasa.gov/, last accessed July 2021) has included an option to download separately NRT active fires from Suomi-NPP and JPSS-1, although it is supposed that in both cases the algorithm described in Schroeder et al. (2014) is applied. At NTC level, a first collection of two L3 products called VNP14ML (750m) and VNP14IMGML (375 m) are currently being produced, although it seems that only Suomi-NPP detections are included. In those products the % confidence level has been replaced by confidence classes (low, nominal, and high).

3.7.3 Land Cover CCI

In the previous Fire_cci phases the source of the Land Cover information was the ESA CCI Land Cover (LC) project. That project produced annual LC maps from 1992 until 2015. More recently, this dataset has been extended within the Copernicus Climate Change Service (C3S) to the years 2016-2019 allowing users to have global and annual updated

LC information at 300 m resolution. The algorithms used in both cases are equal so the consistency of the dataset is ensured (ESA, 2017a). Regarding uncertainty characterisation, there is not such information included in the dataset, although hard work has been done to determine the quality (validation) of the product (ESA, 2017b).

4 Uncertainty characterisation through BA algorithms

Taking into account the current situation of the uncertainty characterisation in the products that are used as input in the Fire_cci project algorithms, a proper characterisation and propagation of the input uncertainties through the algorithms is not feasible. Not only because the different algorithm structures prevent a mathematical propagation through themselves, but also because of the inexistence of such input data. However, an effort is being carried out to analyse how the uncertainty approximations described in the previous section (Section 3) can be used to somehow assess the impact of the algorithms in the final results.

4.1 FireCCISFD20

The FireCCISFD20 product is derived from the MSI sensor aboard two different satellites: Sentinel-2A and Sentinel-2B. Due to a spatial shift between S2A and S2B data, errors were observed to increase when using data from both satellites at the same time. For this reason, the algorithm will generate two independent BA products, each based on images from one satellite, which will be merged in a final step. Each intermediate BA product detects initial burned areas between two dates based on fixed thresholds and proximity to active fires. Then, statistics for MIRBI and NBR2 spectral indices are extracted from these initial BA, which are used to assign the final probability of burn of the whole image, from 0 to 100%. In this way, a probability image is obtained for the image from every single date (Roteta et al., 2019).

Both BA products corresponding to S2A and S2B satellites are fused in the final FireCCISFD20 product. In a temporal series of $\{t_{0,A} , t_{1,B} , t_{2,A} , t_{3,B} , …, t_n\}$ where every element corresponds to a date (with $_A$ and $_B$ standing for S2A and S2B dates), burned pixels in any S2A date ($t_{i,A}$) are assigned a 0% probability if they were not detected as burned in the last day of observation before $t_{i,A}$ ($t_{i-1,B}$) or in the first day of observation after $t_{i,A}$ ($t_{i+1,B}$) in the S2B product. In this way, many commissions due to unmasked clouds and cloud shadows are reduced. Finally, the monthly product is created by assigning to each pixel the date when the highest probability was observed. This probability is generated in two stages. First, s-shape and z-shape sigmoid membership functions applied to temporal differences of MIRBI and NBR2, respectively, depending on their class. In that phase, the main idea is to establish a threshold value dependent on the 'not burned' background and already detected burned areas, using the 90th percentile of not burned background pixels as the minimum value and the 50th percentile of burned pixels as the maximum value of a membership function (Figure 2).
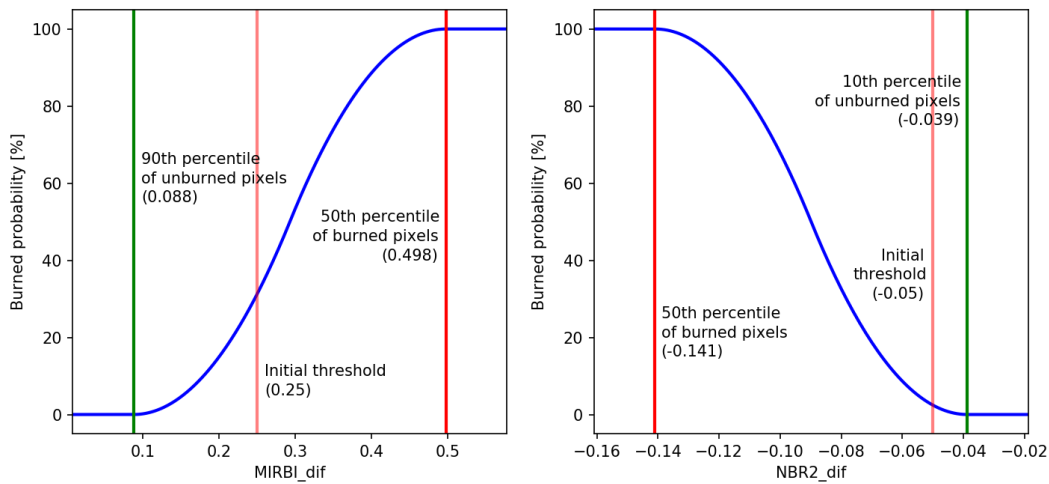
Figure 2. S-shape and Z-shape sigmoid membership functions applied to MIRBI and NBR2. Source: Bastarrika and Roteta (2018)

At the end of this phase, the multiplication of both probability functions ($MIRBI_{burned\_membership}$ and $NBR2_{burned\_membership}$), called Second Stage Probability of Burn (SEPB), is used as the probability of burn to be implemented in the second phase, where these probabilities are rescaled to make it feasible to use the 50% threshold commonly used in binary maps. A lower value should be used in this product due to the following reasons:

- When generating sigmoid curves in the refinement process the percentiles for burned and unburned categories are not symmetric, being the sigmoid curve closer to the burned class than to the unburned one, and thus causing the boundary between both classes to be in a lower probability value.
- The probability value of the SEPB is obtained by multiplying MIRBI and NBR2 probability functions, lowering even more the threshold value.

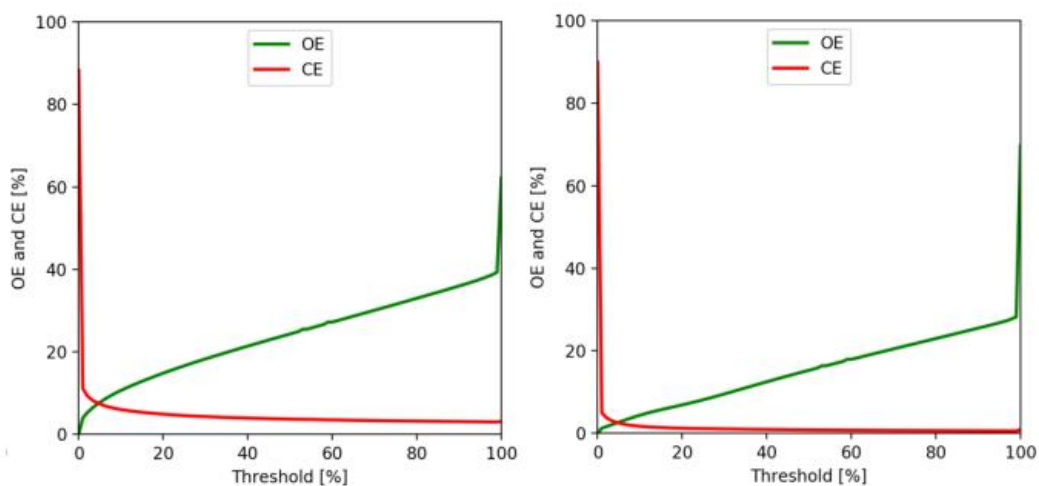**Figure 3. Evolution of omission and commission errors depending on the threshold in two samples. Source: Bastarrika and Roteta (2018)**

In order to decide which was the best threshold, the evolution of both omissions and commissions was analysed depending on the applied threshold. As it is shown in Figure 3, this threshold is around the 5% value. Since this value could be confusing for the product's

users, the probability values were rescaled so that this original 5% value would become the common 50% threshold, following Table 9. Due to the scale of the values being rescaled, the new probability values have discrete values in intervals of 10%. As pixels with a probability of burn below 50% are considered as not burned, only pixels above this threshold are represented in the final SFD product.

Table 2. Lookup Table for the rescaling process

Original probability (%)	Rescaled probability, discrete values (%)
0	0
1	10
2	20
3	30
4	40
5-14	**50**
14-23	60
23-32	70
32-41	80
41-50	90
> =50	100

In the final confidence level product, only probabilities over 50% remain unchanged. A value of 1 is given to observed pixels whose estimated probability of burn was below that threshold. All pixels that have not been observed during the month were given a value of 0 in the confidence level layer. An example of this layer corresponding to the month of June for Area h42v16 is shown in Figure 4.
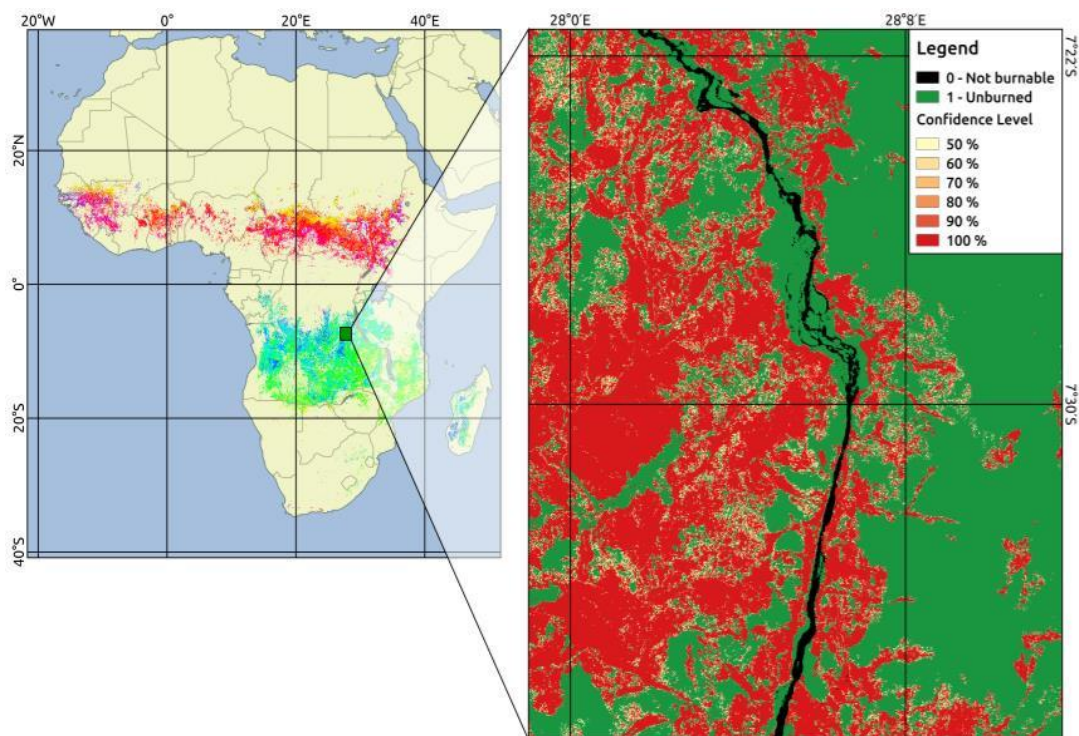

Figure 4. Example of the Confidence Level layer for the h41v19 tile on June 2016. Source: Bastarrika et al, (2019)

4.2 FireCCI51

Regarding this product, which delivers BA within the period 2001-2019. Logistic regression analysis was used in the previous version following a k-Fold validation (k = 10) to estimate BA probability as a surrogate of uncertainty characterization of the BA algorithm. The model was trained from 10 calibration sites, which encompass a wide range of different vegetation and fire patterns, using data from 2008. Four input variables were included in the logistic models: monthly NIR composite (NIR), monthly relative NIR drop ($Rel\Delta NIR$), distance to the nearest BA seed ($distance$), and the number of valid observations in the first 10 post-fire days to the nearest active fire hotspot day (obs) (Lizundia-Loiola et al., 2018). These variables were chosen to be the proxy as they are assumed to be highly linked to the probability of burn. The lower is the NIR value, the higher is the probability of the pixel to be burned, and the lower is the uncertainty. On the other hand, the higher is the relative NIR drop, the higher is the probability of the pixel being burned in the burned classification. It is also assumed that the closer the pixel is to the nearest seed, the higher is its probability of burn. Concerning the latter variable, which can take values from 0 to 10, the higher is the number of observations, the lower is the uncertainty in the observation.

The final coefficients derived from the logistic regression analysis were:

$$p_b = \frac{1}{1 + e^c}$$

(6)

where p_b is the BA probability and $c = -(3.533 - 0.01175 * obs - 0.001996 * NIR + 0.01417 * Rel\Delta NIR - 0.0009282 * distance)$. An example of an area of Northern Australia is shown in Figure 5.
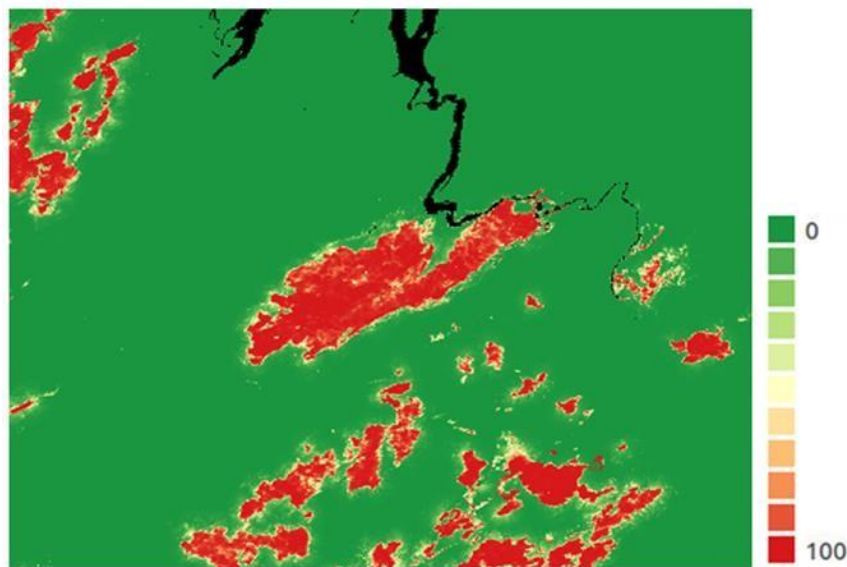
Figure 5. Uncertainty of an area of the tile h30v10 (Australia) for June 2008. The values represent the probability of each pixel being burned, expressed in percentage. Source: Lizundia-loiola et al. (2018)

This method provided a proxy to propagate uncertainty and overcome the challenge of simulating such a complex thresholding algorithm. However, although the layers were assumed to be highly linked to the algorithm components, the logistic regression failed to translate that link in some cases. For instance, the number of observations obs had a coefficient that means the inverse sense of the expected; the lower is the number of

observations, the higher is the probability of burn. This inconsistency is likely to be caused by the multicollinearity character of predictive variables, which has been proved to be inconvenient when using logistic regression (Ranganathan et al., 2017). To cope with these issues the same methodology applied in the case of FireCCIS310 has been adapted to FireCCI51. Nonetheless, the above-mentioned variables are updated to improve the linkage with the algorithm decisions. Instead of using the number of observations in the first 10 post-fire days, it is more convenient to include the number of minima taken from this time window. When the burn probability is high, all the three minima should ideally be detected in that window, but in order to take account of the number of valid observations within the same time frame, the ratio is used. Thus, this variable will be:

$$min_{post} = \frac{N_{min}}{obs} \qquad (7)$$

where N_{min} is the number of minima and obs is the number of valid observations, both in the 10 post-fire days to the nearest hotspot day. The higher the min_{post} , the more probable the burn event is.

The connectivity is also important for this product. However, the most probable burn date was merely generated for the burned class. Another measure of connectivity would be the number of burned neighbours (n_B). The pixels where this value is equal to zero the algorithm classify them as unburned, and the pixels on the edge of burned areas (one or two burned neighbours) are unlikely to be burned. Therefore, the higher is this variable, the more likely is the burn event. In the end, NIR, $Rel\Delta NIR$, $distance$, min_{post} and n_B are used to characterise each pixel. The rest of the process is similar to FireCCIS310 uncertainty estimation using spatiotemporal patterns (Section 4.2).

4.3   FireCCIS310

This product refers to the BA derived from S3 SYN data. The validation dataset generated for 2019 will be used to generate the uncertainty characterization of the FireCCIS310 product. BA algorithm inputs and output over all validation areas are considered to estimate the performance of the algorithm over a large variety of inputs. To reduce the dimensionality of that highly redundant data set and to drastically reduce the computational effort to compute uncertainties, a look-up table (LUT) of representative spatiotemporal signal patterns can be derived from that data set. A reasonable number of distinct spatiotemporal patterns is taken into account from the validation data set. Given the representativeness of validation data, those should be sufficient to describe the full range of typical spatiotemporal patterns in the data. This method was developed by Professor Manuel Campagnolo (personal communication) of the University of Lisbon, who was involved in the previous phase of the project.

The LUT comprises two dimensions, the spatiotemporal patterns and the predictive variables. The latter dimension consists of a number of intermediate processing variables that are able to serve as a proxy to characterize the uncertainty of pixels belonging to each specific spatiotemporal pattern. In the FireCCIS310 the main input variable considered is the NBR2 index, but in the processing, we are mainly interested in the absolute drop of NBR2, which can be calculated based on the difference between the post average and the pre average (eight pre- and eight post-compositing day observations). The rationale behind including this variable is that burned pixels show a remarkable decrease of NBR2 that lasts a considerable period, and therefore, the lower is the value of NBR2 absolute drop, the

more probable is the pixel to be actually burned. Moreover, most of thresholds and operations applied by FireCCIS310 algorithm involve the use of that layer. The separability between averaged NBR2 pre and post is another important parameter as the compositing process relies on it. A high value of separability indicates that is very likely to be due to fire event as a rapid change in the index time series (NBR2 in our case) will yield a separability of large magnitude (Giglio et al., 2009), but to complement this information, the difference between the day of maximum separability and the day of the nearest preliminary seed should be also taken into account. Ideally, this difference should be very small in the case of burn. The last parameter to consider is related to the connectivity of pixels. This layer is important in the sense that neighbouring burned pixels are ideally very close in terms of date of burn. To characterise this concept we consider the temporal texture (Giglio et al., 2009), which can be computed using the standard deviation between the DOY of the pixel under consideration, which is considered as the centre of the series, and its neighbouring pixels (rook's case). Figure 6 illustrate an example of the four variables. It can be noticed that all of them indicate more or less the extent of the burned area.
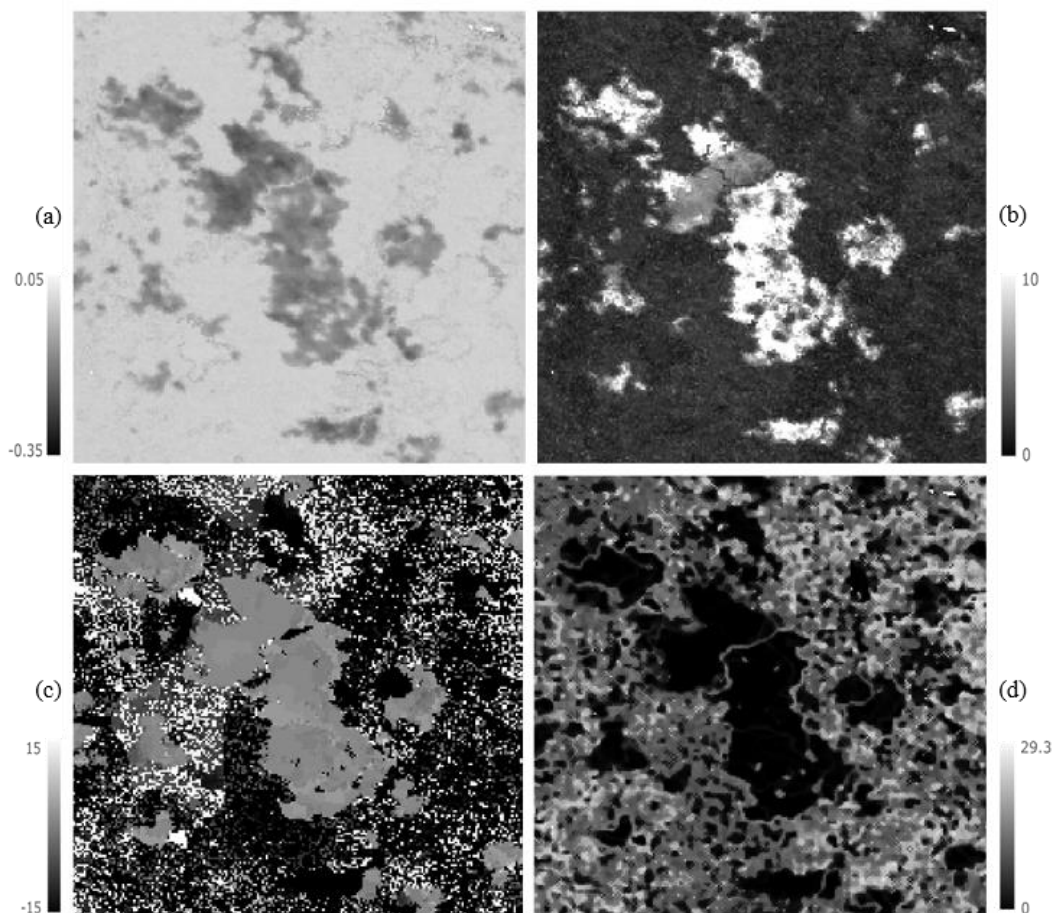
**Figure 6. Example of predictive variables of uncertainty of FireCCIS310: (a) *AbsΔNBR2*;
(b) *MaxSep*; (c) *DoyMaxSep*; (d) *Tex***

In that way, each pixel can be characterised by a monthly vector of four variables: $AbsΔNBR2$, $MaxSep$, $DoyMaxSep$ and Tex. In the next step, the validation process is considered in order to assign another two variables (burned or unburned in reference and algorithm classifications, respectively) to each pixel. Reference data of the year 2019 obtained from the C3S global (2017-2019) dataset of the Burned Area Reference Database (BARD) (Franquesa et al., 2020) is employed. The validation sites of this dataset were selected through stratified random sampling to properly represent the variety of biomes at

the global scale. At the end of this stage, each pixel will be characterised by a set of six variables (Table 3). However, as stated earlier, the objective is to create representative spatiotemporal patterns (P) based on clustering analysis. In order to optimise the computation time, *Mini Batch k-means* classifier of *scikit-learn* library was chosen.

The clustering is based on the four selected variables, and using the BA algorithm outputs and the reference, one can estimate from all N cases that belong to the spatiotemporal pattern P, the four conditional probabilities, as combinations of True/False Positive/Negative. The output of this step is illustrated in Table 3.

Table 3. Example of the matrix that can be obtained from an algorithm

	Variables				Classification	
	$Abs\Delta NBR2$	$MaxSep$	$DoyMaxSep$	Tex	Reference	Algorithm
Pixel 1	X_{11}	X_{12}	X_{13}	X_{14}	B or UB	B or UB
Pixel 2	X_{21}	X_{22}	X_{23}	X_{24}	B or UB	B or UB
Pixel 3	X_{31}	X_{32}	X_{33}	X_{34}	B or UB	B or UB
Pixel 4	X_{41}	X_{42}	X_{43}	X_{44}	B or UB	B or UB
…	…	…	…	…	…	…
Pixel N	X_{N1}	X_{N2}	X_{N3}	X_{N4}	B or UB	B or UB
Spatiotemporal pattern P	X_1	X_2	X_3	X_4	$\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix}$	

Uncertainty of that set of "representative patterns" can be then estimated directly using the confusion matrix. Here, the Sørensen–Dice coefficient (F1 Score), which has been widely used in burned area validation (Belenguer-Plomer et al., 2021; Fornacca et al., 2017; Lizundia-Loiola et al., 2020; Padilla et al., 2014, 2015; Roteta et al., 2019, 2021; Zhang et al., 2021) was employed. The index can be defined as:

$$DC = \frac{2TP}{2TP + FP + FN} \qquad (8)$$

The result is a collection of representative patterns (P_1, P_2,...), and associated "probabilities of burn detection" (pb_1, pb_2,...) that will be saved in a LUT. To compute an uncertainty layer for the whole burned area product, each pixel has to be associated to the most similar representative pattern and labelled with the corresponding probability of burn. Another useful information to be implemented for each pixel, is the uncertainty of the attributed probability. In clustering problems, the most the representative measures of uncertainty is the Euclidian distance to the centre of the nearest cluster, so that the higher values indicate a high uncertainty:

$$u(x) = d_x = \sqrt{\sum_{i=1}^{4} d_{xi}^2} \qquad (9)$$

where d_x is the Euclidian distance to the centre of the nearest cluster (kmeans clusters), and i denotes the predictive variable. Note that these input variables are standardised, and thus we can easily rescale the uncertainty to the probability in order to be considered as a standard deviation using the formula:

$$u_p(x) = \frac{d_x}{4} * p_b(x) \tag{10}$$

where $p_b(x)$ is the probability of burn attributed to the pixel x. It can be also noticed that this standard measure may give unreliable results if the distribution of clusters density is not homogeneous as it does not take into account the proximity to other clusters. In fact, it can happen that two clusters have very similar means and are thus very close in the terms of predictive variables. In this case, the small distance associated with the pixel will indicate that it is close to the centre of the cluster but will not be able to inform about the proximity to the other cluster, which is very close as well. Therefore, the result will indicate a low uncertainty value, which is not actually correct. To correct for that error, a correction factor based on the weighted inverse distance to all clusters is used. This factor can be formulated as:

$$f = \frac{1}{\left[\frac{1/d_x}{\sum_{k=1}^{m} 1/d_k}\right]} \tag{11}$$

where m is the number of clusters. In that way, the relative contribution of very far clusters is negligible as their inverse distances will be very small compared to $1/d_x$, which means that $f \approx 1$. In contrast, the presence of very close clusters will result in a correction factor higher than 1 as the sum of inverse distances will be higher than $1/d_x$.

Finally, the standard uncertainty can be estimated as:

$$u_p(x) = \frac{d_x}{4} * p_b(x) * f \tag{12}$$

In theory, propagating uncertainties from the lower levels of processing is possible using Monte Carlo methods, providing a sound alternative to tracing uncertainty through the algorithm. Given uncertainties in L1 products, Monte Carlo simulations can be performed to quantify the distribution D of predictive variables that classify a pixel as burned/unburned. For S3, the error layer that is provided along with the reflectance in SYN (Section 3.3) may be used as a proxy of the input uncertainty. This would also allow quantifying the stability of the algorithm and its sensitivity to noise in L1 products. The lack of deliverable error layers at L1 SYN at the time of writing that document, as well as the very high computational load of that method for a global product makes this task unfeasible.

4.4 FireCCILT20 and FireCCIS1SA10

Both the FireCCILT20 and FireCCIS1SA10 products are based on random forest models, and therefore, the same uncertainty characterisation and propagation framework can be developed in both cases, although they use different input datasets. Regarding the main

input of the algorithms, AVH09 reflectance (Section 3.4) is used by FireCCILT20 while FireCCIS1SA10 make use of Sentinel 1 SAR data (Section 3.6). Land cover CCI (Section 3.7.3) is used by the two algorithms. Active fire information provided by MODIS (Section 3.7.1) and VIIRS (Section 3.7.2) is also used by the FireCCIS1SA10.

In the case of random forest classifier, quantification of the classification probabilities at pixel level is derived from the proportion of votes for the BA class (classifying trees). Each pixel is attributed a probability distribution, with the probability $p(i)$ of being classified into the class i, given by:

$$p(i) = \frac{k_i}{k} \tag{13}$$

where k_i is the number of trees classifying the pixel as class i, and k is the total number of classification trees in the Random Forests classifier. Here in the case of FireCCILT20, the parameter k was 600, whereas k was 250 for FireCCIS1SA10. However, it should be noted that this uncertainty is just a description of how much one can trust the interpretation of the burned/unburned state of a pixel given the uncertainty of the data.

In the final confidence level layer, the labelling convention is slightly different between the two products as illustrated by Table 4 and Table 5.

Table 4. Confidence level values for FireCCILT20

Attribute	Possible values
Confidence level (CL)	• 0: when the probability is so close to 0 that the classifier rounds the value to it. • 1 to 100: Probability values. When the pixel is closer to 100, that pixel has higher confidence that the pixel is actually burned. This value expresses the uncertainty of the detection for all pixels, even if they are classified as unburned. • -1: when the pixel is not observed in the month. • -2: used for pixels that are not burnable: continuous water, bare land, urban, permanent ice-snow.

Table 5. Confidence level values for FireCCIS1SA10

Attribute	Possible values
Confidence level (CL)	• 0: when the pixel is not observed in the month, or it is not burnable. • 1: value assigned when the pixel was observed, but it was classified as not burned. • 2 to 100: Probability values. The closer to 100, the higher the confidence that the pixel is actually burned.

An example of the confidence level layer of the global FireCCILT20 corresponding to August 1982 is shown in Figure 7.
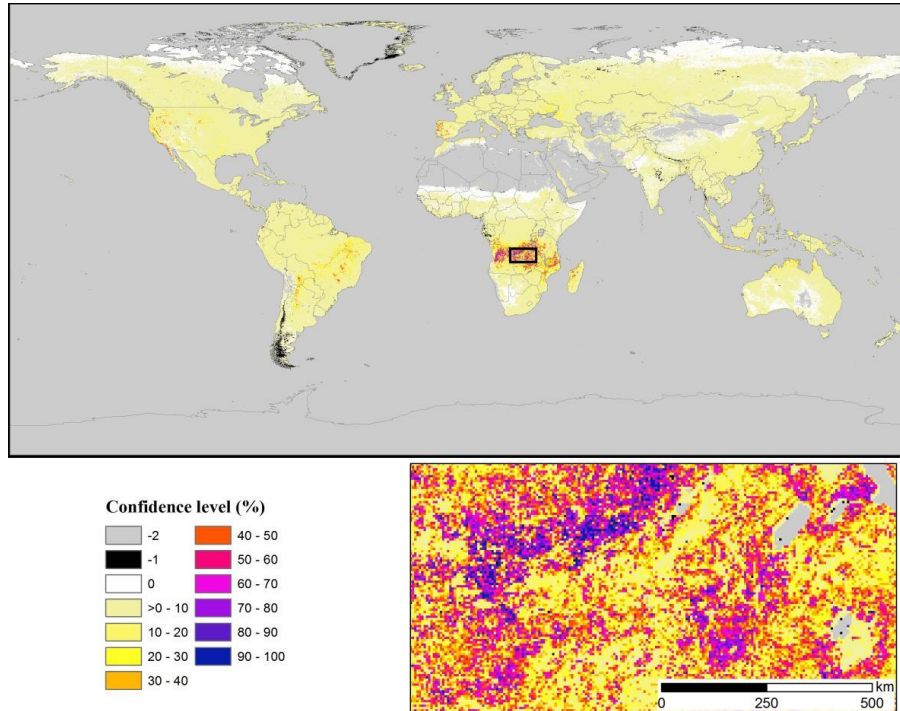
Figure 7. Example of the Confidence Level for FireCCILT20 (August 1982). Source: Otón (2020)

4.5   FireCCIS1S2AF10

FireCCIS1S2AF10 product is based on a deep learning model, more specifically on a Convolutional Neural Network (CNN). The algorithm uses MSI and SAR data (Section 3.5 and 3.6, respectively) as main inputs, in addition to land cover CCI (Section 3.7.3) as well as active fire information provided by MODIS (Section 3.7.1) and VIIRS (Section 3.7.2). Such remote sensing datasets are generally affected by noise and disturbances. Hence, instead of 'real' measurements ($\mathbf{v}$), what is actually fed to the model is a noisy version ($\mathbf{z}$) of such measurements (Loquercio et al., 2020).

The last layer of all CNN-based models is a Softmax layer (Krizhevsky et al., 2012), which is a logistic regression that normalizes an input value ($\mathbf{z}$) into probability values that ranges from 0 to 1:

$$P(\mathbf{z}_i) = \frac{e^{z_i}}{\Sigma_{j=1}^{n} e^{z_j}} \tag{14}$$

where $\mathbf{z}$ is the output of a fully-connected neural network of the class i (i.e. burned or unburned), e is the Euler's number and n is the number of classes (i.e. 2).

Softmax, provides a discrete probability distribution over the burned and unburned classes. This probability value is based on the features extracted in the previous hidden layers of the model, which are affected by convolutional processes that considered the intrinsic uncertainty sources ($\mathbf{z}$). Due to the fact that only two classes (i.e., burned and unburned) were involved in the process, the probability of a given pixel to belong to the burned class (b) is inversely proportional to belong to the unburned one (u):

$$1 = P_b + P_u \tag{15}$$

Hence, such probability distribution may be used as a proxy to estimate the uncertainty. When low intrinsic errors of data are considered, the final probability of a pixel to belong to a given class is lower in comparison to when a high degree of interferences of errors exists (Rottmann & Schubert, 2019; Pascual et al., 2018). The Softmax layer may be considered, therefore, to characterise the uncertainty when mapping BA.

Other authors, however, pointed out that in CNN models the representation of the uncertainty could be poor (Jain et al., 2018). Since this deep learning approach can be considered like "black boxes", there is no way to mathematically propagate the input error through the algorithm. Instead, to somehow improve this representativeness of that error in the final probability, an approach that could be followed is to include the uncertainty layers of each predictive variable in the model. Ideally speaking, the uncertainty of all input variables should be provided in a pixel-based uncertainty layer. When this is not the case, a preliminary uncertainty layer could be created based on the literature (Section 3.4). The aim of those "invented" uncertainties would not be to characterise the actual input uncertainty, but to create a layer that will serve to develop the propagation framework.

To illustrate the mentioned approach, let's consider the NDVI, which is one of the predictive layers of the algorithm. An uncertainty layer corresponding to that variable can be estimated through error propagation framework. Lewis et al. (2018) showed how uncertainty could be generated when ratio-difference transformations are applied. In the case of sums and subtractions, if it is assumed that errors are uncorrelated, the variance of the result is simply equal to the sum of the variances. For the ratio transformation, since it is not possible to directly apply linear propagation, it is common to approximate this by linear terms using a first-order Taylor approximation instead. Mathematical expressions can be found in Lewis et al. (2018). The result of all this process would be an uncertainty layer per each predictive variable. However, the problem with this approach is that it cannot be applied if an actual and proper uncertainty characterisation of the input datasets is not available. Moreover, when the uncertainty is included as a predictive variable the results can be artificially altered. Therefore, the only option is to consider that the probability of burn given by the CNN correctly characterises the uncertainty.

5   Uncertainty characterisation at grid scale

From the point of view of the ESA Fire_cci pixel-level products, there are two relevant layers: the date of the first detection, and the confidence level. Generally speaking, the burned area inside a 0.25° Climate Model Grid (CMG) cell can be determined as the sum of pixels where the first date of detection is between 0 and 366 (both inclusive), multiplied by the area of the pixel. In the previous version of the E3UB, the uncertainty aggregation CMG cell was carried out following the approach developed by University College London that is described in Section 4 of Lewis et al. (2018). This method is preserved for all products except FireCCIS310 and FireCCI51, in which the assumption of independence of cell pixels is unreliable. In this section, we are going to describe both, the previous (hereafter called "Uncertainty using Poisson Binomial distribution") and the new (hereafter called "Expansion of the LPU") methods.

5.1   Uncertainty using Poisson Binomial distribution

This approach can be easily applied also in the current project since it uses the probability of burn generated by the previously mentioned algorithms as input to propagate the uncertainty. This is a number between 0 (absolute certainty that the pixel did not burn in

the time interval considered) to 1 (absolute certainty that the pixel did burn in the time interval considered).

The confidence layer is interpreted as a probability of burn, p_b, (and as a consequence, the probability of not being burned is $1 - p_b$), then this information would need to be scaled up to the CMG, as a form of standard error. There are two common definitions relating to standard error (Weisstein, 2017): (i) the square root of the estimated error variance (standard deviation); (ii) the standard error of a sample of size n is the sample standard deviation divided by $\sqrt{n}$. There is a need to consider then which would be appropriate in this context.

The sample variance σ^2 of a sample set of size n is given by:

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{16}$$

where y_i is sample i and $\bar{y}$ is the sample mean, given by

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \tag{17}$$

The sample terms $\bar{y}$ and σ^2 are random variables, and the expected value of the variance $\hat{\sigma}^2$ is given by

$$\hat{\sigma}^2 = \frac{n}{n-1} \sigma^2 \tag{18}$$

Often, σ^2 is the biased sample variance and $\hat{\sigma}^2$ is the unbiased sample variance. Going back to the definitions of standard error, it can be said that the first definition $\hat{\sigma}_1$ (square root of the estimated error variance) is thus

$$\hat{\sigma}_1 = \frac{n}{n-1} \sigma \tag{19}$$

where σ is the sample standard deviation. Using the second definition (sample standard deviation divided by $\sqrt{n}$) we have that

$$\hat{\sigma}_2 = \frac{1}{\sqrt{n}} \sigma \tag{20}$$

The first definition is more consistent with many uses of standard error in the physical sciences, where it takes the role of an unbiased estimate of the standard deviation of a distribution. If the distribution is assumed normal and y is continuous (or effectively continuous if n is large), then the estimate of the mean ($\bar{y}$) and standard deviation (σ_1) fully define the Probability Distribution Function (PDF) for BA.

The second definition is more directly related to the uncertainty of the mean and is used in the definition of probable error. The standard error of the mean is given by σ_2. Therefore, with more samples (greater n) we can better estimate the mean of the distribution.

In the light of this, we will use

$$\hat{\sigma} = \frac{1}{\sqrt{n-1}} \sum_{i=1}^{n} (y_i - \bar{y})^2 \qquad (21)$$

which is an unbiased estimate of the likely variability in burned area.

Assuming that each pixel has an independent probability of burn p_b, which can be different for every pixel, and then the sum of these independent probabilities is given by a Poisson Binomial distribution. This distribution is only defined over positive integer numbers, and has first and second order statistics given by

$$\overline{N_b} = \sum_{i=1}^{N_p} p_{b,i} \qquad (22)$$

$$\sigma_b^2 = \sum_{i=1}^{N_p} p_{b,i}(1 - p_{b,i}) \qquad (23)$$

In Figure 8, the full PDF derived from a set of samples each characterised by a different p_b is shown. We calculate the PDF as a Poisson binomial, and also calculate the mean and variance using the equations above, and plot the normal approximation to the PDF. For a large number of samples, the skewness of the PDF is very low, and the PDF is acceptably approximated by a Gaussian distribution. This is of importance, as it means that one can parametrize the full PDF of BA using only the mean and the "standard error" (defined as the standard deviation in the discussion above), and in accordance to the product specification.
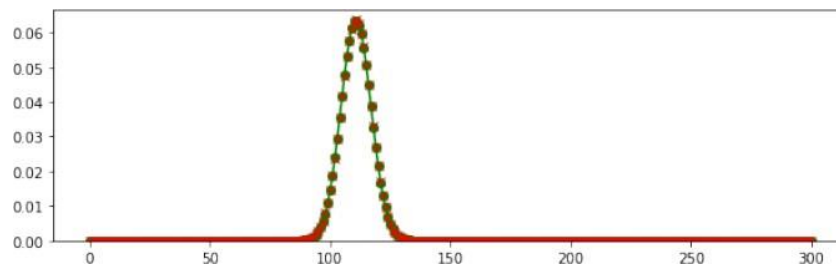
Figure 8. The Poisson binomial PDF (green line) derived from a simulated set of independent samples (300, 100 with probabilities between 0.7 and 0.9, 100 with probabilities between 0.2-0.3 and 100 with probabilities between 0-0.1). A Gaussian approximation (red line) derived from calculating the mean (~110) and standard deviation (~39) is also shown. Skewness was ~0.01. Source: Lewis et al. (2018).

5.2 Expansion of the LPU

In the case of FireCCIS310 and FireCCI51, probabilities at pixel level were obtained from spatiotemporal clusters. Thus, the independence condition assumed to apply Poisson statistical approach is not guaranteed. Instead, the second-order uncertainty layer of each

pixel (standard error of the confidence level) will serve to propagate uncertainty at CMG cells. The aggregated burned area estimation can be approximated to:

$$B = A * \sum_{i=1}^{n} p_{bi} \tag{24}$$

where A is the pixel area, and p_{bi} is the probability of burn of the ith pixel of the CMG cell. Then, the LPU can be applied to this model (Section 2.3). Here the predictive variables are the probabilities of the different pixels inside the CMG cell under consideration.

All sensitivity coefficients are equal to the area of a single pixel, A:

$$C_i = \frac{df}{dp_{bi}} = A \tag{25}$$

In the end, the standard error can be estimated using the formula:

$$\sigma^2(B) = C\, S_x\, C^T \tag{26}$$

where S_x is the covariance matrix of pixel probabilities of burn.

This approach is aligned with recommendations of the GUM and fully linked to the uncertainties at pixel levels. Therefore, it can be more convenient than the previous one if the requirements are satisfied (linear/or approximatively linear model with input uncertainties). In that way, all kinds of error correlations are accounted for and no assumptions needs to be made.

6 References

ACRI-ST IPF Team (2017). Product Data Format Specification - OLCI Level 1 Products, version 2.2. Mission Performance Centre.

ACRI-ST IPF Team (2018). Product Data Format Specification - SLSTR Level 1 Products, version 2.7. Mission Performance Centre.

Anfinsen, S. N., Doulgeris, A. P., & Eltoft, T. (2009). Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery. IEEE Transactions on Geoscience and Remote Sensing, 47(11), 3795-3809.

Ackerman, S., Frey, R., Strabala, K., Liu, Y., Gumley, L., Baum, B., & Menzel, P. (2010). Discriminating clear-sky from cloud with MODIS algorithm theoretical basis document (MOD35). Available at: https://modis.gsfc.nasa.gov/data/atbd/atbd_mod06.pdf

Bastarrika, A., Roteta, E. (2018). ESA CCI ECV Fire Disturbance: D2.1.2 Algorithm Theoretical Basis Document-SFD, version 1.0. Available at: https://www.esa-fire-cci.org/documents.

Bastarrika, A., Roteta, E., Pettinari, M.L., Storm, T. (2019). ESA CCI ECV Fire Disturbance: D3.3.2. Product User Guide – Small Fire Database, version 1.2. Available at: https://www.esa-fire-cci.org/documents.

Belenguer-Plomer M. A., Tanase, M. A., Fernandez-Carrillo, A., & Chuvieco, E. (2018). Temporal backscattering coefficient decorrelation in burned areas. Proc. SPIE 10788, Active and Passive Microwave Remote Sensing for Environmental Monitoring II, 107880T (9 October 2018); https://doi.org/10.1117/12.2325650.

Belenguer-Plomer, M.A., Tanase, M.A., Chuvieco, E., Bovolo, F. (2021). CNN-based burned area mapping using radar and optical data. Remote Sens. Environ. 260, 112468. https://doi.org/10.1016/j.rse.2021.112468.

Belenguer-Plomer, M. A., Tanase, M. A., Fernandez-Carrillo, A., & Chuvieco, E. (2019). Burned area detection and mapping using Sentinel-1 backscatter coefficient and thermal anomalies. Remote Sensing of Environment, 233, 111345.

Brennan, J., Gómez-Dans, J.L., Disney, M., & Lewis, P. (2019). Theoretical uncertainties for global satellite-derived burned area estimates. Biogeosciences, 16, 3147-3164

Bulgin, C.E., Embury, O., & Merchant, C.J. (2016). Sampling uncertainty in gridded sea surface temperature products and Advanced Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC) data. Remote Sensing Of Environment, 177, 287-294

Burnicki, A.C. (2011). Modeling the Probability of Misclassification in a Map of Land Cover Change. Photogrammetric Engineering & Remote Sensing, 77, 39-49.

Chuvieco, E., Pettinari, M.L., Heil, A., & Storm, T. (2017) ESA CCI ECV Fire Disturbance: D1.2 Product Specification Document, version 6.3. Available at: https://www.esa-fire-cci.org/documents

Crosetto, M., Moreno, J.A., & Crippa, B. (2001). Uncertainty propagation in models driven by remotely sensed data. Remote Sensing of Environment, 76, 373-385.

Crosetto, M., & Tarantola, S. (2001). Uncertainty and sensitivity analysis: tools for GIS-based model implementation. International Journal of Geographical Information Systems, 15, 415-437.

El Saleous, N., Vermote, E., Justice, C., Townshend, J., Tucker, C., & Goward, S. (2000). Improvements in the global biospheric record from the Advanced Very High Resolution Radiometer (AVHRR). International Journal of Remote Sensing, 21, 1251-1277.

ESA (2015). Land Cover CCI: Comprehensive Error Characterisation Report - Year 2 version 1.1. Tech. Rep Available at: http://www.esa-landcover-cci.org/?q=webfm_send/124.

ESA (2017a). Land cover CCI product user guide version 3. Tech. Rep Available at: https://www.esa-landcover-cci.org/?q=webfm_send/123.

ESA (2017b). Land cover CCI algorithm theoretical basis document, Part III: classification year 2, version 1.2. Tech. Rep Available at: https://www.esa-landcover-cci.org/?q=webfm_send/139.

EUMESAT (2011). AVHRR Level 1b Product Guide. In: EUM/OPS-EPS/MAN/04/0029 Issue: v3A Date: 21 Jan 2011. Available from: https://www.eumetsat.int/website/ wcm/idc/idcplg?IdcService=GET_FILE&dDocName=PDF_AVHRR_L1B_PRO DUCT_GUIDE&RevisionSelectionMethod=LatestReleased&Rendition=Web

Etxaluze, M., & Smith, D. (2019). Sentinel-3 SLSTR Uncertainties in Level-1 Products Algorithm and Theoretical Basis Document. RAL Space for Copernicus.

Farrance, I., Frenkel, R. (2014). Uncertainty in measurement: A review of monte carlo simulation using microsoft excel for the calculation of uncertainties through functional relationships, including uncertainties in empirically derived constants. Clin. Biochem. Rev. 35, 37–61.

Fornacca, D., Ren, G., Xiao, W. (2017). Performance of Three MODIS fire products (MCD45A1, MCD64A1, MCD14ML), and ESA Fire_CCI in a mountainous area of Northwest Yunnan, China, characterized by frequent small fires. Remote Sens. 9, 1–20. https://doi.org/10.3390/rs9111131.

Franquesa, M., Vanderhoof, M.K., Stavrakoudis, D., Gitas, I.Z., Roteta, E., Padilla, M., Chuvieco, E. (2020). Development of a standard database of reference sites for validating global burned area products. Earth Syst. Sci. Data 12, 3229–3246. https://doi.org/10.5194/essd-12-3229-2020.

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., van der Werf, G.R., Kasibhatla, P., Collatz, G.J., Morton, D.C., & Defries, R. (2010). Assessing variability and long-term trends in burned area by merging multiple satellite fire products. Biogeosciences, 7, 1171-1186.

Giglio, L., Schroeder, W., & Justice, C. O. (2016). The collection 6 MODIS active fire detection algorithm and fire products. Remote Sensing of Environment, 178, 31-41.

Giglio, L., Boschetti, L., Roy, D.P., Humber, M.L., & Justice, C.O., (2018a). The Collection 6 MODIS burned area mapping algorithm and product. Remote Sensing of Environment. 217, 72–85.

Giglio, L., Schroeder, W., Hall, J. V., & Justice, C. O. (2018b). MODIS Collection 6 Active Fire Product User's Guide, Revision B. NASA.

Goncalves, J., & Fernandes, J. C. (2005). Assessment of SRTM-3 DEM in Portugal with topographic map data. In Proceedings of the EARSeL Workshop 3D-Remote Sensing, 10-11.

Gorroño, J., Fomferra, N., Peters, M., Gascon, F., Underwood, C.I., Fox, N.P., Kirches, G., & Brockmann, C. (2017). A radiometric uncertainty tool for the Sentinel 2 mission, Remote Sensing, 9, 178.

GUM (2008), JCGM 100:2008 Guide to the expression of Uncertainty in Measurement (JCGM) p 100 (www.bipm.org)

GUM-101 (2008), JCGM 101:2008, Evaluation of measurement data — Supplement 1 to the "Guide to the expression of uncertainty in measurement" — Propagation of distributions using a Monte Carlo method (www.bipm.org)

Holben, B.N., Kaufman, Y.J., & Kendall, J.D. (1990). NOAA-11 AVHRR visible and near-IR inflight calibration. International Journal of Remote Sensing, 11, 1511-1519.

Hollmann, R., Merchant, C.J., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., Chuvieco, E., Defourny, P., De Leeuw, G., Forsberg, R., Holzer-Popp, T., Paul, F., Sandven, S., Sathyendranath, S., Van Roozendael, M., Wagner, W. (2013). The ESA climate change initiative: Satellite data records for essential climate variables. Bull. Am. Meteorol. Soc. 94, 1541–1552. https://doi.org/10.1175/BAMS-D-11-00254.1

Hunt, S., & Nieke, J. (2016). A Radiometric Uncertainty Tool for OLCI. In, Proceedings Living Planet Symposium, 9-13.

Jain, A., Fandango, A., & Kapoor, A. (2018). TensorFlow Machine Learning Projects: Build 13 real-world projects with advanced numerical computations using the Python ecosystem. Packt Publishing Ltd.

Kalogirou, V., Ferrazzoli, P., Della Vecchia, A., & Foumelis, M. (2014). On the SAR backscatter of burned forests: A model-based study in C-band, over burned pine canopies. IEEE Transactions on Geoscience and Remote Sensing, 52(10), 6205-6215.

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).

Kurum, M. (2015). C-band SAR backscatter evaluation of 2008 Gallipoli forest fire. IEEE Geoscience and Remote Sensing Letters, 12(5), 1091-1095.

Lewis, P., Gomez Dans, J., Brennan, J., & Chernetskiy, M. (2018) ESA CCI ECV Fire Disturbance: D3.1. Comprehensive Error Characterization Report, version 2.0. Available at: https://www.esa-fire-cci.org/documents

Lizundia-Loiola, J., Otón, G., Ramo, R., & Chuvieco, E. (2020). A spatio-temporal active-fire clustering approach for global burned area mapping at 250 m from MODIS data. Remote Sensing of Environment, 236, 111493.

Lizundia-Loiola, J., Pettinari, M.L., Chuvieco, E., Storm, T., & Gómez-Dans, J. (2018) ESA CCI ECV Fire Disturbance: D2.1.3 Algorithm Theoretical Basis Document-MODIS, version 2.0. Available at: https://www.esa-fire-cci.org/documents

Loquercio, A., Segu, M., & Scaramuzza, D. (2020). A general framework for uncertainty estimation in deep learning. IEEE Robotics and Automation Letters, 5(2), 3153-3160.

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., & Gascon, F. (2017). Sen2Cor for Sentinel-2. In Image and Signal Processing for Remote Sensing XXIII p. 1042704. International Society for Optics and Photonics.

Merchant, C., Paul, F., Popp, T., Ablain, M., Bontemps, S., Defourny, P., Hollmann, R., Lavergne, T., Laeng, A., de Leeuw, G., Mittaz, J., Poulsen, C., Povey, A., Reuter, M., Sathyendranath, S., Sandven, S., Sofeiva, V., Wagner, W. (2017). Uncertainty information in climate data records from Earth observation. Earth Syst. Sci. Data Discuss. 1–28. https://doi.org/10.5194/essd-2017-16

Mittaz, J., Merchant, C.J., Woolliams, E.R. (2019). Applying principles of metrology to historical Earth observations from satellites. Metrologia 56. https://doi.org/10.1088/1681-7575/ab1705

North, P. & Heckel, A. (2010) Sentinel-3 Optical products and algorithm definition. SYN Algorithm Theoretical Basis Document.

Otón, G., Ramo, R., Lizundia-Loiola, J., & Chuvieco, E. (2019). Global Detection of Long-Term (1982–2017) Burned Area with AVHRR-LTDR Data. Remote Sensing, 11, 2079.

Otón, G. (2020). ESA CCI ECV Fire Disturbance: D4.2.2 Product User Guide – AVHRR-LTDR, version 1.0. Available at: https://www.esa-fire-cci.org/documents

Padilla, M. & Chuvieco, E. (2014). ESA CCI ECV Fire Disturbance: D1.2.3 Comprehensive Error Characterisation Report, version 2.1. Available at: https://www.esa-fire-cci.org/documents

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 Sens. Environ. 160, 114–121. https://doi.org/http://dx.doi.org/10.1016/j.rse.2014.01.008.

Pascual, G., Seguí, S., & Vitrià, J. (2018). Uncertainty Gated Network for Land Cover Segmentation. In CVPR Workshops (pp. 276-279).

Pettinari, M.L., Chuvieco, E., Lizundia-Loiola, J., & Tanase, M. (2019) ESA CCI ECV Fire Disturbance: D1.2 Algorithm Development Plan, version 1.1. Available at: https://www.esa-fire-cci.org/documents

Pedelty, J., Devadiga, S., Masuoka, E., Brown, M., Pinzon, J., Tucker, C., Vermote, E., Prince, S., Nagol, J., Justice, C., Roy, D., Ju, J., Schaaf, C., Liu, J., Privette, J., & Pinheiro, A. (2007). Generating a long-term land data record from the AVHRR and MODIS instruments. In, Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International, 1021-1025.

Quegan, S., & Yu, J. J. (2001). Filtering of multichannel SAR images. IEEE Transactions on Geoscience and Remote Sensing, 39(11), 2373-2379.

Ranganathan, P., Pramesh, C.S., Aggarwal, R. (2017). Common pitfalls in statistical analysis: Logistic regression. Perspect. Clin. Res. 8, 148. https://doi.org/10.4103/PICR.PICR_87_17.

Robel, J., Graumann, A., Kidwell, K., Aleman, R., Ruff, I., Muckle, B., & Kleespies, T. (2014). NOAA KLM User's Guide with NOAA-N, N Prime, and MetOp Supplements. National Oceanic and Atmospheric Administration, Washington.–

2530 p. Available at: https://www1.ncdc.noaa.gov/pub/data/satellite/publications/ podguides/N-15_thru_N-9/pdf/0.0_NOAA_KLM_Users_Guide.pdf

Rochman, D., van der Marck, S.C., Koning, A.J., Sjöstrand, H. & Zwermann, W. (2014). Uncertainty Propagation with Fast Monte Carlo Techniques, Nuclear Data Sheets, 118, 367-369.

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 Sens. Environ. 222, 1–17. https://doi.org/10.1016/j.rse.2018.12.011

Roteta, E., Bastarrika, A., Franquesa, M., Chuvieco, E. (2021). Landsat and sentinel-2 based burned area mapping tools in google earth engine. Remote Sens. 13, 1–30. https://doi.org/10.3390/rs13040816

Rottmann, M., & Schubert, M. (2019). Uncertainty measures and prediction quality rating for the semantic segmentation of nested multi resolution street scene images. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (9 pp.). Available at: http://openaccess.thecvf.com/content_CVPRW_ 2019/papers/SAIAD/Rottmann_Uncertainty_Measures_and_Prediction_Quality_ Rating_for_the_Semantic_Segmentation_CVPRW_2019_paper.pdf

S3 Team (2011). Level 1 Algorithms Theoretical baseline Document – Part 2: optical products [SY-24] Level 1C ATBD. Thales Alenia Space.

Santoro, M., Cartus, O., Quegan, S. (2019). Biomass CCI: End-to-End ECV Uncertainty Budget, year 1, version 1.0. Available at: https://climate.esa.int/en/projects/biomass/key-documents.

Schroeder, W., Oliva, P., Giglio, L., & Csiszar, I. A. (2014). The New VIIRS 375m active fire detection data product: algorithm description and initial assessment, Remote Sensing of Environment, 143, 85-96.

Schwerdt, M., Schmidt, K., Tous Ramon, N., Klenk, P., Yague-Martinez, N., Prats-Iraola, P., Zink, M., & Geudtner, D. (2017). Independent system calibration of Sentinel-1B. Remote Sensing, 9(6), 511.

Smith, J.H., Stehman, S.V., Wickham, J.D., & Yang, L. (2003). Effects of landscape characteristics on land-cover class accuracy. Remote Sensing of Environment, 84, 342-349.

Stoffelen, A. (1998). Toward the true near-surface wind speed: Error modelling and calibration using triple collocation, J. Geophys. Res.- Oceans, 103, 7755–7766.

Trishchenko, A.P., Fedosejevs, G., Li, Z., & Cihlar, J. (2002). Trends and uncertainties in thermal calibration of AVHRR radiometers onboard NOAA-9 to NOAA-16. Journal of Geophysical Research: Atmospheres, 107 (D24), 4788.

Uprety, S., & Cao, C. (2011). Using the Dome C site to characterize AVHRR near-infrared channel for consistent radiometric calibration. Earth Observing Systems XVI, 81531Y: International Society for optics and Photonics.

van Oort, P.A.J., Bregt, A.K., de Bruin, S., de Wit, A.J.W., & Stein, A. (2011). Spatial variability in classification accuracy of agricultural crops in the Dutch national land-cover database. International Journal of Geographical Information Science, 18, 611-626.

Vermote, E. F., & Saleous, N. (2006). Operational atmospheric correction of MODIS visible to middle infrared land surface data in the case of an infinite Lambertian target. In Earth science satellite remote sensing (pp. 123-153). Springer, Berlin, Heidelberg.

Vermote, E. F., & Kotchenova, S. (2008). Atmospheric correction for the monitoring of land surfaces. Journal of Geophysical Research: Atmospheres, 113(D23).

Vermote, E., Roger, J., & Ray, J. (2015). MODIS Surface Reflectance User's Guide-Collection 6. Tech. Rep. Version 1.4, NASA GSFC Terrestrial Information Systems Laboratory, MODIS Land Surface Reflectance Science Computing Facility: Greenbelt, USA.

Vickers, H., Blanco, P., Duguay, C. (2020). Lakes CCI: End-to-End ECV Uncertainty Budget, version 1.0. Available at: https://climate.esa.int/en/projects/lakes/key-documents-lakes.

Weisstein, Eric W (2017). "Standard Error." From MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/StandardError.html

Xiong, J., Toller, G., Sun, J., Wenny, B., Angal, A., & Barnes, W. (2013). MODIS level 1B algorithm theoretical basis document, version 4. NASA MODIS characterization support team, Washington, DC.

Zhang, P., Ban, Y., Nascetti, A. (2021). Learning U-Net without forgetting for near real-time wildfire monitoring by the fusion of SAR and optical time series. Remote Sens. Environ. 261, 112467. https://doi.org/10.1016/j.rse.2021.112467

Annex 1 Acronyms and abbreviations

AC	Atmospheric Correction
ADC	Analog-to-Digital Converter
ADF	Auxiliary Data Files
ADP	Algorithm Development Plan
AERONET	Aerosol Robotic Network
AOD	Aerosol Optical Depth
AOT	Atmospheric Optical Thickness
ASCII	American Standard Code for Information Interchange
ATBD	Algorithm Theoretical Basis Document
AVH09	AVHRR Surface Reflectance Product of Land Long Term Data Record project
AVHRR	Advanced Very High Resolution Radiometer
BA	Burned area
BC	Brockmann Consult GmbH
BOA	Bottom-of-Atmosphere
BRDF	Bidirectional reflectance distribution function
BT	Brightness Temperature
C3S	Copernicus Climate Change Service
CCD	Charge-Coupled Device
CCI	Climate Change Initiative
CDR	Climate Data Record
CMG	Climate Model Grid
CNN	Convolutional Neural Networks
CRG	Climate Research Group
dB	decibel
E3UB	End-to-End ECV Uncertainty Budget
ECV	Essential Climate Variable
ENL	Equivalent Number of Looks
Envisat	Environmental Satellite
EO	Earth Observation
ERS	European Remote Sensing satellite
ESA	European Space Agency
EUMESAT	European Organisation for the Exploitation of Meteorological Satellites
EV	Earth View
FIDUCEO	Fidelity and Uncertainty in Climate data records from Earth Observation
FireCCI51	Fire_cci MODIS version 5.1
FireCCILT10	Fire_cci AVHRR-LTDR version 1.0
FireCCILT20	Fire_cci AVHRR-LTDR version 2.0
FireCCIS1S2AF10	Fire_cci Sentinel-1 & Sentinel-2 test sites in AFrica version 1.0
FireCCIS310	Fire_cci Sentinel-3 version 1.0
FireCCISFD20	Fire_cci Small Fire Database version 2.0
FIRMS	Fire Information for Resource Management System

GAC	Global Area Coverage
GFED	Global Fire Emissions Database
GRD	Ground Range Detected
GUM	Guide to the Expression of Uncertainty in Measurement
IFOV	Instantaneous Field of View
IR	Infrared
ISP	Instrument Source Packet
IW	Interpherometric Wide
JCGM	Joint Committee for Guides in Metrology
JPSS	Joint Polar Satellite System
K	Kelvin
L0	Level 0
L1	Level 1
L2	Level 2
L3	Level 3
L4	Level 4
LC	Land Cover
LPU	Law of Propagation of Uncertainty
LTDR	Long Term Data Record
LUT	Look-Up-Table
LWL	Lakes Water Level
MapnoiS3	Sentinel-3 Noise Mapping python tool
MCD14ML	MODIS Global Monthly Fire Location Product
MCD64	MODIS Collection 5 and 6 Burned Area product using the Giglio et al. (2009) and Giglio et al. (2018a) algorithm, respectively
MCST	MODIS Characterization and Support Team
MIRBI	Mid-Infrared Burn Index
MOD09	MODIS Terra Surface Reflectance 5-minute L2 Swath
MOD09GA	MODIS Terra Surface Reflectance Daily L2G Global 500 m and 1 km
MOD09GQ	MODIS Terra Surface Reflectance Daily L2G Global 250 m
MOD14/MYD14	MODIS Terra/Aqua active fire and other thermal anomalies product
MOD35	MODIS Cloud Mask product
MODIS	Moderate Resolution Imaging Spectroradiometer
MODLAND	MODIS Land Team
MSI	MultiSpectral Instrument
NASA	National Aeronautics and Space Administration
NBR2	Normalized Burn Ratio 2
NetCDF	Network Common Data Form
NIR	Near-InfraRed
NOAA	National Oceanic and Atmosphere Administration
NPP	National Polar-orbiting Partnership
NRT	Near Real Time
NTC	Non-Time Critical

OLCI	Ocean and Land Colour Instrument
OZA	OLCI Zenith Angle
PDF	Probability Distribution Function
PDGS	Payload Data Ground Segment
PRT	Platinum Resistance Thermometer
PSD	product Specification Document
PUG	Product user Guide
QA	Quality Assurance
QA4EO	Quality Assurance Framework for Earth Observation
QI	Quality Indicators
S1	Sentinel-1
S2	Sentinel-2
S2-RUT	Sentinel-2 radiometric uncertainty Tool
S3	Sentinel-3
SAR	Synthetic Aperture Radar
SCM	Scene Classification Map
Sen2Cor	Sentinel 2 Level 2A product generation and formatting processor
SLC	Single Look Complex
SLSTR	Sea and land Surface Temperature Radiometer
SNAP	Sentinels Application Platform
SRF	Spectral Response Function
SRTM DEM	Shuttle Radar Topography Mission Digital Elevation Model
SST	Sea Surface Temperature
Suomi-NPP	Suomi National Polar-orbiting Partnership satellite
SWIR	Short-Wave InfraRed
SY_1_MISR	Internal Synergy Level 1 product
SY_2_SYN	Synergy Surface Directional Reflectance product
SYN	Synergy
TC	Triple Collocation
TCI	True Color Image
TOA	Top-of-Atmosphere
TOPSAR	Terrain Observation with Progressive Scans SAR
VIIRS	Visible Infrared Imaging Radiometer Suite
VISCAL	Visible Calibration unit
VNP14ML	VIIRS Global Monthly Fire Location Product at 750 m
VNPIMG14ML	VIIRS Global Monthly Fire Location Product at 375 m
WV	Water Vapour