https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

Open Access Earth System Science Data Discussions

Development of a standard database of reference sites for validating global burned area products

Magí Franquesa[1], Melanie K. Vanderhoof[2], Renata Libonati[3,4], Julia A. Rodrigues[3], Alberto W. Setzer[5], Dimitris Stavrakoudis[6], Ioannis Z. Gitas[6], Ekhi Roteta[7], Marc Padilla[8], Emilio Chuvieco[1]

[1]Environmental Remote Sensing Research Group, Department of Geology, Geography and the Environment, Universidad de Alcalá, Calle Colegios 2, Alcalá de Henares, 28801, Spain
[2]Geological Survey, Geosciences and Environmental Change Science Center, P.O. Box 25046, DFC, MS980, Denver, CO 80225, United States
[3]Departamento de Meteorologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 21941-916, Brazil
[4]Centro de Estudos Florestais, Instituto Superior de Agronomia, Universidade de Lisboa, Lisboa, 1349-017, Portugal
[5] Centro de Previsão de Tempo e Estudos Climáticos/Instituto Nacional de Pesquisas Espaciais, Programa de Monitoramento de Queimada por Satélites, 12227-010, São José dos Campos, SP, Brazil
[6]Laboratory of Forest Management and Remote Sensing, School of Forestry and Natural Environment, Aristotle University of Thessaloniki, P.O. Box 248, GR-54124, Greece
[7]Department of Mining and Metallurgical Engineering and Materials Science, School of Engineering of Vitoria-Gasteiz, University of the Basque Country UPV/EHU, Nieves Cano 12, Vitoria-Gasteiz, 01006, Spain
[8]Centre for Landscape& Climate Research, Department of Geography, University of Leicester, Leicester LEI 17RH, United Kingdom

Correspondence to: Magí Franquesa (magin.franquesa@uah.es)

Abstract. Over the past two decades, several global burned area products have been produced and released to the public. However, the accuracy assessment of such products largely depends on the availability of reliable reference data that currently does not exist on a global scale or whose production requires high level dedication of project resources. The important lack of reference data for the validation of burned area products is addressed in this paper. We provide the first Burned Area Reference Database (BARD) that was created by compiling existing reference burned area datasets from different international projects. The Database contains a total of 2769 reference burned area files derived from Landsat or Sentinel-2 imagery. All reference files have been checked for internal quality and are freely provided by the authors. To ensure database consistency, all files were transformed to a common format and were properly documented by following metadata standards. This should help future users of this database to read and convert the files to their own preferred formats or projections. The database is freely available at: https://doi.org/10.21950/BBQQU7 (Franquesa et al., 2020).

1 Introduction

Validation is defined by the Committee on Earth Observing Satellites Working Group on Calibration and Validation (CEOS-WGCV) as "the process of assessing, by independent means, the quality of the data products derived from the system outputs" (CEOS-WGCV, 2012). Validation helps in evaluating the utility and limitations of using any remote sensing (RS) product,

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

particularly on whether user accuracy requirements are met. For this reason, validation should be part of any RS project, even though it requires additional effort and cost that is not aimed at improving accuracy but rather to measure it. Validation implies comparing our results to reference data, assumed to represent the actual conditions of the target variable at the satellite overpass time. In the case of global studies, it is very difficult to generate reference data for the wide variety of planetary conditions, thereby complicating validation. Some of the global variables (e.g. temperature and surface radiation) can be validated from ground sensors networks, such as weather stations, buoys or Aerosol Robotic NETwork (AERONET) sensors. Other variables are more difficult to validate, as they require generating global reference data that is based on higher-resolution sensors than those used to obtain the global product. This is the case of land cover or burned area products, which require first designing a sample strategy using statistically valid protocols and then extracting from the selected sites the reference polygons to be compared with the global datasets. Despite the time and effort required to derive reference datasets, accuracy assessment is a critical part of any global RS project and making these reference datasets publicly available will facilitate product comparison and lower the burden of validating future product efforts.

Several global burned area (BA) products have been produced in the last two decades, providing an estimation of fire activity worldwide (Chuvieco et al., 2019). The first global BA product, Global Burned Area (GBA2000) was based on daily VEGETATION (VGT) images acquired for 2000 and was generated at coarse resolution by the Joint Research Centre of the European Union (Grégoire et al., 2003). The same year, the European Space Agency developed the GLOBSCAR BA product, derived from daytime ERS-2 (European Remote Sensing Satellite) ATSR-2 (Along Track Scanning Radiometer) data (Simon et al., 2004). Both BA products, GBA2000 and GLOBSCAR BA, had a nominal pixel size of 1 km^2. Other 1 km resolution global BA products released by European projects include the L3JRC (Tansey et al., 2008) covering the period from 2000 to 2007; GlobCarbon (Plummer et al., 2006), produced from 1998 to 2007; and the Copernicus GIO_GL1_BA products. All these products were derived from VGT images (in GlobCarbon, ATSR images were used as well). More recently, the FireCCI project (https://esa-fire-cci.org, last access: 25 March 2020), part of the European Space Agency's (ESA) Climate Change Initiative, has generated three global BA products, based on MERIS (FireCCI41: Alonso-Canas and Chuvieco, 2015) and MODIS data (FireCCI50: Chuvieco et al., 2018;FireCCI51: Lizundia-Loiola et al., 2020), the latter at 250m spatial resolution. NASA released in 2010 the MCD45A1 product derived from 500 m MODIS imagery (Roy et al., 2008), which has now been superseded by MCD64A1 (Giglio et al., 2009;2018).

These global BA products have been validated by comparing them with reference data generated from medium resolution sensors (such as those on board the Landsat, SPOT or Sentinel-2 missions). These reference files were typically derived from multitemporal pairs of images to properly date the validation period. Early validation exercises were usually based on small samples of reference sites without probabilistic meaning (i.e. direct sampling), based on data availability and expert knowledge to ensure main wildfire conditions were included in the sample (Tansey et al., 2004;Roy et al., 2005). Roy and Boschetti (2009), for instance, reported validation results for the MCD45A1 product using a set of 11 Landsat scenes distributed across southern Africa. Chuvieco et al. (2008) validated a regional product for Latin America using 19 Landsat scenes and 9 China–Brazil Earth Resources Satellite (CBERS) scenes. The MCD64A1 Collection 5 was not formally validated, and the most recent

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

Open Access Earth System Science Data Discussions

MCD64A1 Collection 6 products were first validated using a set of 108 Landsat scenes distributed across a wide range of fire-affected ecosystems (Giglio et al., 2018). A recent study has provided a full spatio-temporal validation of the MCD64A1 product, using Landsat-8 Operational Land Imager (OLI) images, but only for a single year (Boschetti et al., 2019). Previous statistical validation of NASA and FireCCI products have been conducted by Padilla et al. (2014;2015) using a set of 130 images for a single year (2008) and by Chuvieco et al.(2018) using a multitemporal reference dataset of 12 years. Other projects covering large areas have been developed in the USA using Landsat data (Vanderhoof et al., 2017) and Africa using Sentinel-2 Multispectral Instrument (MSI) images (Roteta et al., 2019). In both cases, reference datasets were created based on independent interpretation of burn perimeters, controlled by visual inspection.

The main bottleneck for validating global BA products or global BA algorithms is the generation of reference BA datasets. To facilitate the activity of BA algorithm developers, this paper aims to deliver to the scientific community a set of reference BA perimeters that can be used as reference data for validation of BA algorithms. These validation files were compiled from different international projects and years, therefore facilitating the test of BA algorithms in a wide range of ground conditions. The paper presents the methods that were used to generate the BA reference perimeters and then to transform all the files to a common standard format and file structure.

2 Methods

2.1 Selection of validation sites

High quality reference data generation is an expensive and time-consuming task, which constrains the total amount of validation sites that can be established in any validation exercise. For this reason, sampling design is critical to make the most of the resources available and ensure the highest precision of accuracy estimates given the available resources to generate reference data. Padilla et al. (2014;2015) implemented a stratified random sampling that allowed for global BA accuracy inferences for the first time. Boschetti et al. (2016) improved the sampling by specifically including the temporal dimension at the sampling units. More recently, Padilla et al. (2017) presented a first approach to efficiently stratify the population and allocate sample across strata. A probability sampling design, for example a stratified random sampling, allows for unbiased population estimates, and is commonly defined by spatial and temporal dimensions. The spatial dimension of sampling units are usually defined by the Thiessen scene areas (TSAs) constructed by Cohen et al. (2010) and Kennedy et al. (2010) specifically for use with Landsat WRS-2 frames (Fig. 1a). The key advantage of TSAs is that they provide non-overlapping Landsat-like frames, which allow for a convenient computation of unbiased estimators (Gallego, 2005). The temporal dimension of sample units is defined by the acquisition dates of consecutive images.

Sample units are then stratified to properly represent the variety of conditions that affect the accuracy of BA products. This stratification is usually based on (a) major Olson biomes (Olson et al., 2001) (Fig.1b) and (b) the BA extent provided by a considered reliable global BA product or active fire detections, assigning each sample unit to a high or low BA strata based on

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

a threshold that can be specifically adapted to each biome stratum as in Padilla et al. (2017) or simply set as the 20th quantile of the cumulative distribution of active fire counts as in Boschetti et al. (2016;2019).

2.2 Reference data generation methods

Following the recommendations of the CEOS Calibration/Validation group, all the reference burn perimeters were derived from multi-temporal comparison of medium resolution satellite imagery (Landsat TM/ETM+/OLI or Sentinel-2 MSI). Burned patches included in the reference files are only those that occurred in between the two satellite images used to generate the files (Fig.2). The procedure to obtain those burn patches is diverse, depending on the dataset, but all include a semi-automatic procedure (e.g. Bastarrika et al., 2011) and then a visual inspection to confirm that the detected perimeters were actually burned areas. Optionally, the classification can be overwritten by polygons digitized manually. In several cases, this visual inspection was confirmed by another interpreter to double check the quality. Parts of the scene that cannot be observed or interpreted, either by clouds or by sensor problems (i.e. Scan Line Corrector (SLC)-off problems of ETM+) in the pre- or post-image were classified as no-data. This was done to make sure that only areas with reliable data were included in the validation process.

2.3 Data specifications

Existing validation datasets from different global and regional projects were compiled in the Burned Area Reference Database (BARD). Each dataset is organised in three folders with associated files including: (a) 'metadata', which contains a .csv file containing the file name of all the reference files included in such dataset, along with additional information like the temporal length (days) or the total number of images used for each reference file; (b) 'regions', which contains a shapefile containing all the sample sites (TSAs or Sentinel-2 tiles) covered by the dataset; and (c) 'shapefiles', containing the validation reference shapefiles ordered by year.

Reference files are released as Esri shapefile (*.shp) format in UTM/WGS84 projection. The name of the files is defined as follows: 'Project_RD_ppprrr_yyyymmdd_yyyymmdd' (e.g. FireCCI_RD_164069_20160514_20160709'), where:

Project = Project in which the reference data were generated.

RD = stands for Reference Data.

ppprrr = refers to the Landsat Worldwide Reference System (WRS) path (ppp) and row (rrr) of the scene. For collections where Sentinel-2 was used instead of Landsat images, ppprrr refers to the Sentinel-2 tile (e.g. FireCCI_RD_T28PET_ 20160111_20160311').

yyyymmdd (year, month, day). The first date corresponds to the pre-fire date, which is the date of the first image used for BA detection; the second one refers to the post-fire date, which is the date of the last image used for generating the reference fire perimeters.

The following attribute fields are included in the shapefiles (Table 1):

- category:
    - 1: Burned area. This category includes all polygons detected as burned

- 2: No-Data. This category includes all polygons that could not be interpreted or were not observed by the sensor, either by clouds and/or cloud shadows, topographic shadows, smoke, or sensor errors (for instance, those caused by SLC-off problems of ETM+).
- 3: Unburned. This category includes all polygons observed as not burned within the limits of the area covered by the image.

- preDate: Acquisition date of the image taken before the occurrence of the fire: yyyy-mm-dd (year, month, day).
- postDate: Acquisition date of the image taken after the fire: yyyy-mm-dd (year, month, day).
- preImg and postImg: The pre- and post-fire Landsat scene identifier (e.g. 'LC80260422013124LGN01'). For reference files based on S2 images, the datastrip ID is used instead. (e.g. 'S2A_OPER_MSI_L1C_TL_SGS__20160420T171415_A004324_T28PEB_N02.01').
- path: The Worldwide Reference System-2 (WRS-2) path of the Landsat scene. For reference files based on S2, the tile number was used.
- row: The row of the Landsat scene. For reference files based on S2, the tile number was used.
- year: The year of the validation dataset.
- area: Area in square meters (m^2) calculated on the WGS84/UTM Cartesian plane.

2.4 Reference datasets

2.4.1 FireCCI global (2008)

The FireCCI global 2008 reference dataset (Padilla et al., 2014) was created using a stratified random sampling design. Two levels of spatial stratification were used to select the spatial units based on TSAs derived from the Landsat World Reference System II (WRS-II). Spatial units were first stratified across seven aggregated Olson biomes (Olson et al., 2001). Each biome was stratified into high and low BA extent based on the Global Fire Emissions Database (GFED) Version 3 (Giglio et al., 2009;2010). Landsat-5 TM (n=127) and Landsat-7 ETM+ (n=131) images were used to retrieve BA perimeters. The complete scene was used for Landsat-5 TM images, whereas only the center of Landsat-7 ETM+ scenes were interpreted in order to avoid data SLC gaps. BA perimeters were derived using a semi-automatic algorithm developed by Bastarrika et al. (2011), where high burn severity pixels were selected to train core burned area, and adjacent lower burn severity pixels were selected to apply a region-growing algorithm to the previously identified burned areas.

The FireCCI global 2008 dataset includes 129 reference data files, derived from image pairs, for the year 2008. The temporal length of reference files varies between 8 and 144 days: 64% of image pairs were separated by 32 days or less, 28% between 32 and 100 days and 8% more than 100 days with a maximum time gap between the pre- and post-image of 144 days. The location and temporal length of the reference files is shown in Fig. 3. This reference dataset is compliant with CEOS-LPVS Stage 3 (https://lpvs.gsfc.nasa.gov/, last access: 25 March 2020).

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

2.4.2 FireCCI global (2003-2014)

This global dataset covers a period of twelve years, from 2003 to 2014 (Padilla et al., 2018). The reference files were generated from consecutive Landsat images separated by 8-16 days for each selected TSA and year. A total of 2358 images from Landsat-5 TM (n=585), Landsat-7 ETM+ (n=1564) and Landsat-8 OLI (n= 209) satellite sensors were used to retrieve BA perimeters. The sampling units were selected following a stratified random sampling design. The total population of sample units was defined spatially by TSAs and temporally by the dates of Landsat images available, filtering out those with a cloud cover greater than 30%. For each calendar year, the sample units were stratified by Olson biomes (Olson et al., 2001) and BA based on MCD64A1 (Giglio et al., 2009). The threshold used to assign the high/low BA strata to each sample unit was defined separately for each year and biome. Once the strata were defined by year-biome-BA, a set of 100 sampling units were selected for each calendar year applying a proportional allocation according to Eq. (1):

$$n_h = N_h \overline{BA}_h \quad (1)$$

Where n_h is the sample size to be selected in stratum h, N_h is the stratum size and $\overline{BA}_h$ the BA mean in stratum h. Finally, a spatial subset window of 30 x 20 km located at the center of the images was applied for interpretation and BA reference data retrieval. The reference perimeters were extracted from a dedicated Random Forest algorithm, trained for each sampling site and output maps were visually inspected by two interpreters (Padilla et al., 2018).

The FireCCI global (2003-2014) dataset includes 1200 reference data files from 722 different TSAs and twelve years, from 2003 to 2014. The temporal length of reference files varies between 8 and 16 days. The location and total number of reference files in each TSA is shown in Fig. 4. This reference dataset is compliant with CEOS-LPVS Stage 3.

2.4.3 FireCCI Africa (2016)

The FireCCI Africa reference dataset consists in consecutive 8-16 days BA maps and was generated for the year 2016 from Landsat imagery (Padilla et al., 2018). The sampling was designed with sampling units long in its time dimension, hereinafter referred to as long units as opposed to the 8-16 days' short units mentioned in the previous section. Reference data over long units allows for long temporal overlaps between validation and product data, and among other analysis it allows to assess the effect of product temporal errors on the accuracy estimates. Reference maps at long units consist simply in the concatenation of consecutive 8-16 days' maps (Fig. 5). The sampling design was similar to that for the FireCCI global (2003-2014) dataset, as mentioned in the previous section. The only difference was on the sample size, 50 units instead of 100 units per year. Note that each unit here is much larger, as it consists on multiple image pairs. Two reference perimeters datasets are released: (a) Reference data at short units level, 1052 files with 8-16 days BA maps and (b) Reference data at long units level, 50 files. The temporal length covered by each long unit varies from 24 to 256 days (Fig. 6b): 18% of the long units cover a temporal length below 50 days, 34% between 50 and 100 days and 48% are above 100 days. The location, number of short units and temporal length of the long units' reference files are shown in Fig. 6. This reference dataset is compliant with CEOS-LPVS Stage 3.

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

2.4.4 FireCCI Africa S2 (2016)

The FireCCI Africa S2 BA reference dataset was created to perform an initial validation assessment of the Small Fire Database Fire_cci v1.1 product (FireCCISFD11) produced for the year 2016 for the whole Sub-Saharan Africa (Roteta et al., 2019). Reference data files were generated from the comparison of two Sentinel-2 MSI images at 20 m resolution per reference site. A systematic sampling was used to select 52 validation sites based on Sentinel-2 tiles (100 x 100 km) over Sub-Saharan Africa. Burned areas were mapped with the BAMS methodology, which is a semi-automated algorithm (Bastarrika et al., 2014). In short, training polygons for the burned category were defined in each tile, and burned seeds were detected. Then, burned regions were grown out from these seeds until the whole burned patches were covered. The results were visually analyzed to determine the accuracy of the classification and new training polygons were defined if needed. This was done consecutively until all burned areas were mapped and no commission or omission errors were visually detected. Finally, if there was noise created by unmasked clouds and cloud shadows, it was edited and removed manually.

The temporal length of the reference files varies between 10 and 120 days: 86% of the pairs of images were separated by less than 50 days and 14% by more than 50 days with a maximum time lapse of 120 days. The location and temporal length of the reference files are shown in Fig. 7. This reference dataset is compliant with CEOS-LPVS Stage 1.

2.4.5 BrFLAS Brazil (2015)

The BrFLAS Brazil BA reference files were generated by a joined initiative by the Laboratory for Environmental Satellite Applications (LASA) from the Federal University of Rio de Janeiro (UFRJ) and the National Institute of Space Research (INPE) under the scope of the Brazilian Fire-Land-Atmosphere System (BrFLAS) Project (http://idlcc.fc.ul.pt/brflas/index.html, last access: 25 March 2020). The BrFLAS Brazil (2015) dataset includes 84 reference data files for the year 2015 (Rodrigues et al., 2019) covering the 77% of the Cerrado Brazilian biome's surface. The dataset was derived from images mapped every 16 days at a spatial resolution of 30 m using multispectral images from the OLI sensor aboard the Landsat-8 satellite. The BA perimeters were systematically generated using a classification method based on the Normalized Difference Vegetation Index (NDVI) and Normalized Burn Ratio Long-shortwave infrared variation (NBRL) indices, and on the difference of those indices between the post- and pre-images (Melchiori et al., 2014). These burned scars automatically generated were then subject to an independent analysis and visual photo interpretation, including a series of quality control procedures for removing data of reduced accuracy to ensure consistency among all the burn scar samples. Images with cloud coverage greater than 10% were discarded for the analysis and only images from June to October were interpreted to reduce cloud contamination and rainfall episodes. A detailed visual analysis of cloud distribution in pre- and post-fire images was performed to avoid commission errors. Finally, all the mapped burned scars retrieved from the 16 days' consecutive pair of images were merged for each path/row Landsat-8 scene and the corresponding unburned region was added to build the final reference data file. Additional processing details can be found in Rodrigues et al. (2019).

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

The temporal length of reference files varies between 16 and 80 days. The location and temporal length of the reference files as well as the number of images used in each reference site are shown in Fig. 8. The reference dataset's stage of validation should be considered as 1 according to the CEOS-LPVS classification.

2.4.6 BAECV CONUS (1988-2013)

The BAECV reference dataset extends across the conterminous United States (CONUS) and was generated to perform the validation of the Landsat Burned Area Essential Climate Variable (BAECV) product (Hawbaker et al., 2017). The sampling design was adapted from the methods used by the ESACCI FireCCI project. Existing FireCCI validation TSAs (n=9) within CONUS were augmented with an additional 19 TSAs for a total of 28 TSAs. The TSAs were stratified across the major Olson biomes (Olson et al., 2001) including (1) temperate forest, (2) Mediterranean forest, (3) temperate grassland and savannah, (4) tropical and subtropical grasslands and savannah and (5) xeric/desert shrub. TSAs selected within each biome were meant to represent high and low burned areas as specified by the Global Fire Emissions Database (GFED) version 3. A systematic sampling was applied to select 6 validation years spaced out in 5-year increments (2013, 2008, 2003, 1998, 1993 and 1988). A total of 336 images from Landsat-5 TM (n=269), Landsat-7 ETM+ (n=10) and Landsat-8 OLI (n= 56) were used to derive the BA extent. Landsat reference images were limited to those with a geometric Root Mean Square Error (RMSE) < 10 m, <20% cloud cover, and available as a L1T Surface Reflectance product. Time lapse between images was not limited to 16 days and only two (pre and post-fire image) were used to retrieve BA perimeters for each path/row. The pre- and post-fire image pairs did not specifically represent a probability sample within a year but were designed to target changes incurred over the peak fire season. Peak fire season was determined using the distribution of total burned area by month as derived from the MCD45 burned area product (2001-2015). The FMask from the Landsat surface reflectance product was applied to mask out clouds, cloud shadows, snow and open water (Zhu and Woodcock, 2014). For Landsat-7 ETM+ images, SLC off pixels were masked out. The low-, medium- and high-intensity development classes (i.e. urban areas) were masked out using the National Land Cover Database (NLCD) (Homer et al., 2015) to reduce spectral confusion between burned areas and impervious surfaces.

Burned area (post-fire – pre-fire) maps were generated using BAMS (Bastarrika et al., 2014). The Normalized Burn Ratio (NBR), Mid-infrared Burned Index (MIRBI), Global Environmental Monitoring Index (GEMI) and Normalized Difference Vegetation Index (NDVI) were calculated for the pre- and post-fire images and utilized in a supervised classification. The algorithm was trained on manually selected polygons containing (1) clearly burned pixels and (2) spectrally similar but less distinct burned pixels. The algorithm applied a region-growing function between the two types of training polygons, while cut-off values for each variable were extracted from the training polygons. Each classified burned area was then manually edited. When available, the analysts utilized ancillary datasets (e.g. Monitoring Trends in Burn Severity (MTBS), MODIS active fire points, MODIS burned area, aerial imagery) to improve the confidence in their selection of training pixels and manual edits. To maximize the accuracy of the reference dataset, each image pair was classified into burned area extent and visually evaluated and edited independently by three different analysts. A pixel was then classified as burned if it was identified as burned by two of the three analysts. Fires with patch sizes of less than 4.05 ha (45 pixels) were removed from the reference data to be comparable with the minimum mapping unit of the BAECV product. Additional processing details can be found in Vanderhoof et al. (2017).

The BAECV CONUS (1988-2013) dataset includes 168 reference data files from 28 Landsat path/rows and six years (1988, 1993, 1998, 2003, 2008, 2013). The temporal length of reference files varies between 16 and 288 days: 37% of pairs of images were separated by less than 50 days, 35% between 50 and 100 days, and 28% by more than 100 days with a maximum time lapse between the pre and post-image of 288 days. Location of reference sites based on TSAs is shown in Fig. 9. The reference dataset's stage of validation should be considered as 3 according to the CEOS-LPVS classification.

2.4.7 NOFFi Greece (2016-2018)

The reference data were obtained using the perimeters produced by the National Observatory of Forest Fires (NOFFi) (http://epadap.web.auth.gr, last access: 25 March 2020) and, specifically, its Object-based Burned Area Mapping (OBAM) service, implemented by the Laboratory of Forest Management and Remote Sensing (FMRS) of the Aristotle University of Thessaloniki. NOFFi-OBAM is an on-demand service, meaning that it is activated after large wildfires events and under explicit requests by the local forest offices. It relies solely on Sentinel-2 imagery and it is employed only for fires within Greece. NOFFi-OBAM's algorithm is designed to map fire perimeters and follows a supervised learning approach using a postfire Sentinel-2 (Level-1C) image, although a prefire image is also used for photo-interpretation purposes. The methodology applied to retrieve the fire perimeters is fully described in Tompoulidou et al. (2016).

The NOFFi-OBAM fire perimeters were used as basis for creating the reference data for the NOFFi Greece reference dataset considering the burned area mappings of years 2016, 2017 and 2018. For each Sentinel-2 tile ID (e.g. T34SDH) in which there were available fire perimeters, the whole time-series of images were visually checked and the date range for the reference file creation was defined from the first pre-fire image to the last post-fire image. Small fires within the specific time series that were not mapped from the NOFFi-OBAM service were explicitly digitized. Since NOFFi-OBAM only serves Greece, areas outside Greece's official land boundaries were masked out and classified as unobserved surfaces (category = 2). Some burned scars in overlapping border tiles were mapped by using images from those neighboring tiles only if the postfire image used for the mapping was inside the formers tile ID's time span. For example, the file 'NOFFi_RD_T34SGH_20160710_20160802.shp', includes polygons with preImg/postImg from T35SCK. This can be identified from the preImg, postImg, tile columns of the file. Cloudy images were discarded from the analysis to avoid unmapped areas within the mapped region.

The NOFFi Greece dataset includes 34 reference data files from 27 different Sentinel-2 tiles. The temporal length of reference files varies between 5 and 132 days. The location and temporal length of the reference files as well as the number of images used in each reference site is shown Fig. 10. The reference dataset's stage of validation should be considered as 1 according to the CEOS-LPVS classification.

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

3 Data availability

The Burned Area Reference Database compiled in this effort is freely available on the e-cienciaDatos repository (https://doi.org/10.21950/BBQQU7 (Franquesa et al., 2020)). All burned area reference data files have been visually checked, reprojected and reformatted to provide a uniform set of attributes and metadata descriptions to maximize the ease with which these reference files can be used to evaluate global burned area products. A summary of the data included in each dataset is described in Table 2. Reference shapefiles and metadata files can be downloaded grouped by the datasets described in this publication: FireCCI global (2008), FireCCI global (2003-2014), FireCCI Africa (2016), FireCCI Africa S2 (2016), BrFLAS Brazil (2015), BAECV CONUS (1988-2013) and NOFFi Greece (2016-2018). The Burned Area Reference Database will be expanded with new reference files that are being produced in the FireCCI project and we encourage future contributions from the scientific community.

4 Conclusions

The Burned Area Reference Database is the first publicly available database that compiles and standardizes previously generated validation reference data. Datasets from different international projects are released, comprising a total of 2769 shapefiles representing reference burned area data generated from approximately 4800 Landsat and Sentinel-2 images. Thus, this BA reference database and futures updates come to replace the lack of an extensive global and regional, multitemporal validation dataset (Humber et al., 2019) and, certainly, can serve as a valuable source for validation of existing and forthcoming BA algorithms.

5 Author contributions

MF and EC have written the first draft of the manuscript. MF has coordinated the manuscript production and prepared the figures, standardized the reference files and organized the BARD database, and managed its publication on the e-cienciaDatos repository. MV provided the BAECV CONUS (1988-2013) dataset. RL, JR and AS provided the BrFLAS Brazil (2015) fire perimeters. DS and IZG provided the NOFFi Greece (2016-2018) dataset. ER provided the FireCCI Africa S2 (2016) dataset, and MP provided the rest of the FireCCI datasets. EC, as the Science Leader of the FireCCI project, managed the overall execution of the project and suggested the preparation of the present article. All the authors have contributed to the writing and reviewing of the manuscript, and agreed on the final version.

6 Competing interests

The authors declare that they have no conflict of interest.

7 Acknowledgements.

This research has been funded by the FireCCI project (contract no 4000126706/19/I-NB) which is part of the ESA Climate Change Initiative. We thank Joshua J. Picotte (USGS Earth Resources Observation and Science (EROS) Center, USA) and M. Lucrecia Pettinari (University of Alcalá, Spain) for their comments and suggestions that help to improve the manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Alonso-Canas, I., and Chuvieco, E.: Global burned area mapping from ENVISAT-MERIS and MODIS active fire data, Remote Sensing of Environment, 163, 140-152, https://doi.org/10.1016/j.rse.2015.03.011, 2015.

Bastarrika, A., Chuvieco, E., and Martín, M. P.: Mapping burned areas from Landsat TM/ETM+ data with a two-phase algorithm: Balancing omission and commission errors, Remote Sensing of Environment, 115, 1003-1012, https://doi.org/10.1016/j.rse.2010.12.005, 2011.

Bastarrika, A., Alvarado, M., Artano, K., Martinez, M., Mesanza-Moraza, A., Leyre, T., Ramo, R., and Chuvieco, E.: BAMS: A Tool for Supervised Burned Area Mapping Using Landsat Data, Remote Sensing, 6, 12360-12380, https://doi.org/10.3390/rs61212360, 2014.

Boschetti, L., Stehman, S. V., and Roy, D. P.: A stratified random sampling design in space and time for regional to global scale burned area product validation, Remote Sensing of Environment, 186, 465-478, https://doi.org/10.1016/j.rse.2016.09.016, 2016.

Boschetti, L., Roy, D. P., Giglio, L., Huang, H., Zubkova, M., and Humber, M. L.: Global validation of the collection 6 MODIS burned area product, Remote Sensing of Environment, 235, 111490, https://doi.org/10.1016/j.rse.2019.111490, 2019.

Working Group on Calibration and Validation - Land Product Validation Subgroup, 2012.

Chuvieco, E., Opazo, S., Sione, W., Del Valle, H., Anaya, J., Di Bella, C., Cruz, I., Manzo, L., López, G., Mari, N., González-Alonso, F., Morelli, F., Setzer, A., Csiszar, I., Kanpandegi, J. A., Bastarrika, A., and Libonati, R.: Global burned-land estimation in Latin America using MODIS composite data, Ecol Appl, 18, 64-79, 10.1890/06-2148.1, 2008.

Chuvieco, E., Lizundia-Loiola, J., Pettinari, M. L., Ramo, R., Padilla, M., Tansey, K., Mouillot, F., Laurent, P., Storm, T., Heil, A., and Plummer, S.: Generation and analysis of a new global burned area product based on MODIS 250 m reflectance bands and thermal anomalies, Earth Syst. Sci. Data, 10, 2015-2031, 10.5194/essd-10-2015-2018, 2018.

Chuvieco, E., Mouillot, F., van der Werf, G. R., San Miguel, J., Tanasse, M., Koutsias, N., García, M., Yebra, M., Padilla, M., Gitas, I., Heil, A., Hawbaker, T. J., and Giglio, L.: Historical background and current developments for mapping burned area from satellite Earth observation, Remote Sensing of Environment, 225, 45-64, https://doi.org/10.1016/j.rse.2019.02.013, 2019.

Cohen, W. B., Yang, Z., and Kennedy, R.: Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync—Tools for calibration and validation, Remote Sensing of Environment, 114, 2911-2924, 2010.

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

Franquesa, M., Vanderhoof, M. K., Libonati, R., Rodrigues, J. A., Setzer, A. W., Stavrakoudis, D., Gitas, I., Roteta, E., Padilla, M., and Chuvieco, E.: BARD: a global and regional validation burned area database, doi:10.21950/BBQQU7, 2020.

Gallego, F. J.: Stratified sampling of satellite images with a systematic grid of points, ISPRS Journal of Photogrammetry and Remote Sensing, 59, 369-376, https://doi.org/10.1016/j.isprsjprs.2005.10.001, 2005.

Giglio, L., Loboda, T., Roy, D. P., Quayle, B., and Justice, C. O.: An active-fire based burned area mapping algorithm for the MODIS sensor, Remote Sensing of Environment, 113, 408-420, https://doi.org/10.1016/j.rse.2008.10.006, 2009.

Giglio, L., Randerson, J. T., van der Werf, G. R., Kasibhatla, P. S., Collatz, G. J., Morton, D. C., and DeFries, R. S.: Assessing variability and long-term trends in burned area by merging multiple satellite fire products, Biogeosciences, 7, 1171-1186, https://doi.org/10.5194/bg-7-1171-2010, 2010.

Giglio, L., Boschetti, L., Roy, D. P., Humber, M. L., and Justice, C. O.: The Collection 6 MODIS burned area mapping algorithm and product, Remote Sensing of Environment, 217, 72-85, https://doi.org/10.1016/j.rse.2018.08.005, 2018.

Grégoire, J. M., Tansey, K., and Silva, J.: The GBA2000 initiative: developing a global burnt area database from SPOT-VEGETATION imagery, Int. J. Remote Sens., 24, 1369–1376, 10.1080/0143116021000044850, 2003.

Hawbaker, T. J., Vanderhoof, M. K., Beal, Y. J., Takacs, J. D., Schmidt, G. L., Falgout, J. T., Williams, B., Fairaux, N. M., Caldwell, M. K., Picotte, J. J., Howard, S. M., Stitt, S., and Dwyer, J. L.: Mapping burned areas using dense time-series of Landsat data, Remote Sensing of Environment, 198, 504-522, 10.1016/j.rse.2017.06.027, 2017.

Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N., Wickham, J., and Megown, K.: Completion of the 2011 National Land Cover Database for the Conterminous United States - Representing a Decade of Land Cover Change Information, Photogrammetric Engineering and Remote Sensing, 81, 346-354, doi:10.14358/pers.81.5.345, 2015.

Humber, M. L., Boschetti, L., Giglio, L., and Justice, C. O.: Spatial and temporal intercomparison of four global burned area products, Int. J. Digit. Earth, 12, 460-484, 10.1080/17538947.2018.1433727, 2019.

Kennedy, R. E., Yang, Z., and Cohen, W. B.: Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms, Remote Sensing of Environment, 114, 2897-2910, 2010.

Lizundia-Loiola, J., Otón, G., Ramo, R., and Chuvieco, E.: A spatio-temporal active-fire clustering approach for global burned area mapping at 250 m from MODIS data, Remote Sensing of Environment, 236, 111493, https://doi.org/10.1016/j.rse.2019.111493, 2020.

Melchiori, A. E., Setzer, A. W., Morelli, F., Libonati, R., Candido, P., and Jesus, S.: A Landsat-TM/OLI algorithm for burned areas in the Brazilian Cerrado: preliminary results, Advances in Forest Fire Research, 23-30, http://dx.doi.org/10.14195/978-989-26-0884-6_143, 2014.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'Amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., and Kassem, K. R.: Terrestrial Ecoregions of the World: A New Map of Life on EarthA new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity, BioScience, 51, 933-938, https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2, 2001.

Padilla, M., Stehman, S. V., and Chuvieco, E.: Validation of the 2008 MODIS-MCD45 global burned area product using stratified random sampling, Remote Sensing of Environment, 144, 187-196, doi:10.1016/j.rse.2014.01.008, 2014.

Padilla, M., Stehman, S. V., Ramo, R., Corti, D., Hantson, S., Oliva, P., Alonso-Canas, I., Bradley, A. V., Tansey, K., Mota, B., Pereira, J. M., and Chuvieco, E.: Comparing the accuracies of remote sensing global burned area products using stratified random sampling and estimation, Remote Sensing of Environment, 160, 114-121, https://doi.org/10.1016/j.rse.2015.01.005, 2015.

Padilla, M., Olofsson, P., Stehman, S. V., Tansey, K., and Chuvieco, E.: Stratification and sample allocation for reference burned area data, Remote Sensing of Environment, 203, 240-255, https://doi.org/10.1016/j.rse.2017.06.041, 2017.

Padilla, M., Wheeler, J., and Tansey, K.: ESA CCI ECV Fire Disturbance: D4.1.1.Product Validation Report, version 2.1, Available at: https://www.esa-fire-cci.org/documents, 2018.

Plummer, S., Arino, O., Simon, M., and Steffen, W.: Establishing A Earth Observation Product Service For The Terrestrial Carbon Community: The Globcarbon Initiative, Mitigation and Adaptation Strategies for Global Change, 11, 97-111, 10.1007/s11027-006-1012-8, 2006.

Rodrigues, J. A., Libonati, R., Pereira, A. A., Nogueira, J. M. P., Santos, F. L. M., Peres, L. F., Santa Rosa, A., Schroeder, W., Pereira, J. M. C., Giglio, L., Trigo, I. F., and Setzer, A. W.: How well do global burned area products represent fire patterns in the Brazilian Savannas biome? An accuracy assessment of the MCD64 collections, International Journal of Applied Earth Observation and Geoinformation, 78, 318-331, https://doi.org/10.1016/j.jag.2019.02.010, 2019.

Roteta, E., Bastarrika, A., Padilla, M., Storm, T., and Chuvieco, E.: Development of a Sentinel-2 burned area algorithm: Generation of a small fire database for sub-Saharan Africa, Remote Sensing of Environment, 222, 1-17, https://doi.org/10.1016/j.rse.2018.12.011, 2019.

Roy, D. P., Frost, P. G. H., Justice, C. O., Landmann, T., Le Roux, J. L., Gumbo, K., Makungwa, S., Dunham, K., Du Toit, R., Mhwandagara, K., Zacarias, A., Tacheba, B., Dube, O. P., Pereira, J. M. C., Mushove, P., Morisette, J. T., Santhana Vannan, S. K., and Davies, D.: The Southern Africa Fire Network (SAFNet) regional burned-area product-validation protocol, Int. J. Remote Sens., 26, 4265-4292, 10.1080/01431160500113096, 2005.

Roy, D. P., Boschetti, L., Justice, C. O., and Ju, J.: The collection 5 MODIS burned area product — Global evaluation by comparison with the MODIS active fire product, Remote Sensing of Environment, 112, 3690-3707, https://doi.org/10.1016/j.rse.2008.05.013, 2008.

Roy, D. P., and Boschetti, L.: Southern Africa validation of the MODIS, L3JRC, and GlobCarbon burned-area products, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 47, 1032-1044, 2009.

Simon, M., Plummer, S., Fierens, F., Hoelzemann, J. J., and Arino, O.: Burnt area detection at global scale using ATSR-2: The GLOBSCAR products and their qualification, Journal of Geophysical Research: Atmospheres, 109, https://doi.org/10.1029/2003JD003622, 2004.

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

Tansey, K., Grégoire, J. M., Stroppiana, D., Sousa, A., Silva, J., Pereira, J. M., Boschetti, L., Maggi, M., Brivio, P. A., and Fraser, R.: Vegetation burning in the year 2000: Global burned area estimates from SPOT VEGETATION data, Journal of Geophysical Research: Atmospheres, 109, 2004.

Tansey, K., Grégoire, J. M., Defourny, P., Leigh, R., Pekel, J. F., Bogaert, E., and Bartholomé, E.: A new, global, multi-annual (2000–2007) burnt area product at 1 km resolution, Geophysical Research Letters, 35, 1-6, 10.1029/2007gl031567, 2008.

Tompoulidou, M., Stefanidou, A., Grigoriadis, D., Dragozi, E., Stavrakoudis, D., and Gitas, I.: The Greek National Observatory of Forest Fires (NOFFi), Fourth International Conference on Remote Sensing and Geoinformation of the Environment, SPIE, 2016.

Vanderhoof, M. K., Fairaux, N., Beal, Y.-J. G., and Hawbaker, T. J.: Validation of the USGS Landsat Burned Area Essential Climate Variable (BAECV) across the conterminous United States, Remote Sensing of Environment, 198, 393-406, https://doi.org/10.1016/j.rse.2017.06.025, 2017.

Zhu, Z., and Woodcock, C. E.: Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change, Remote Sensing of Environment, 152, 217-234, doi:10.1016/j.rse.2014.06.012, 2014.

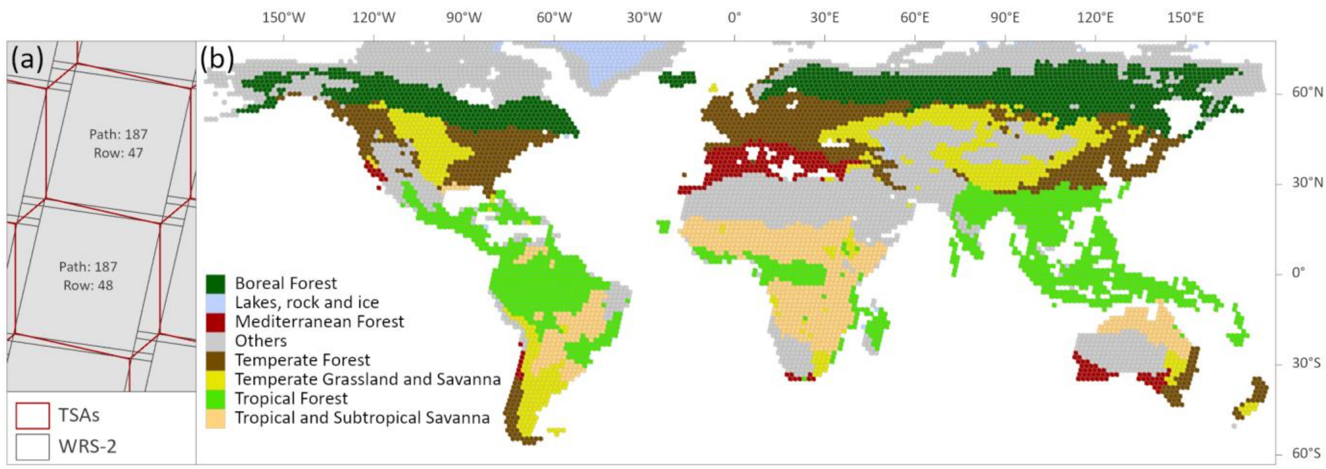
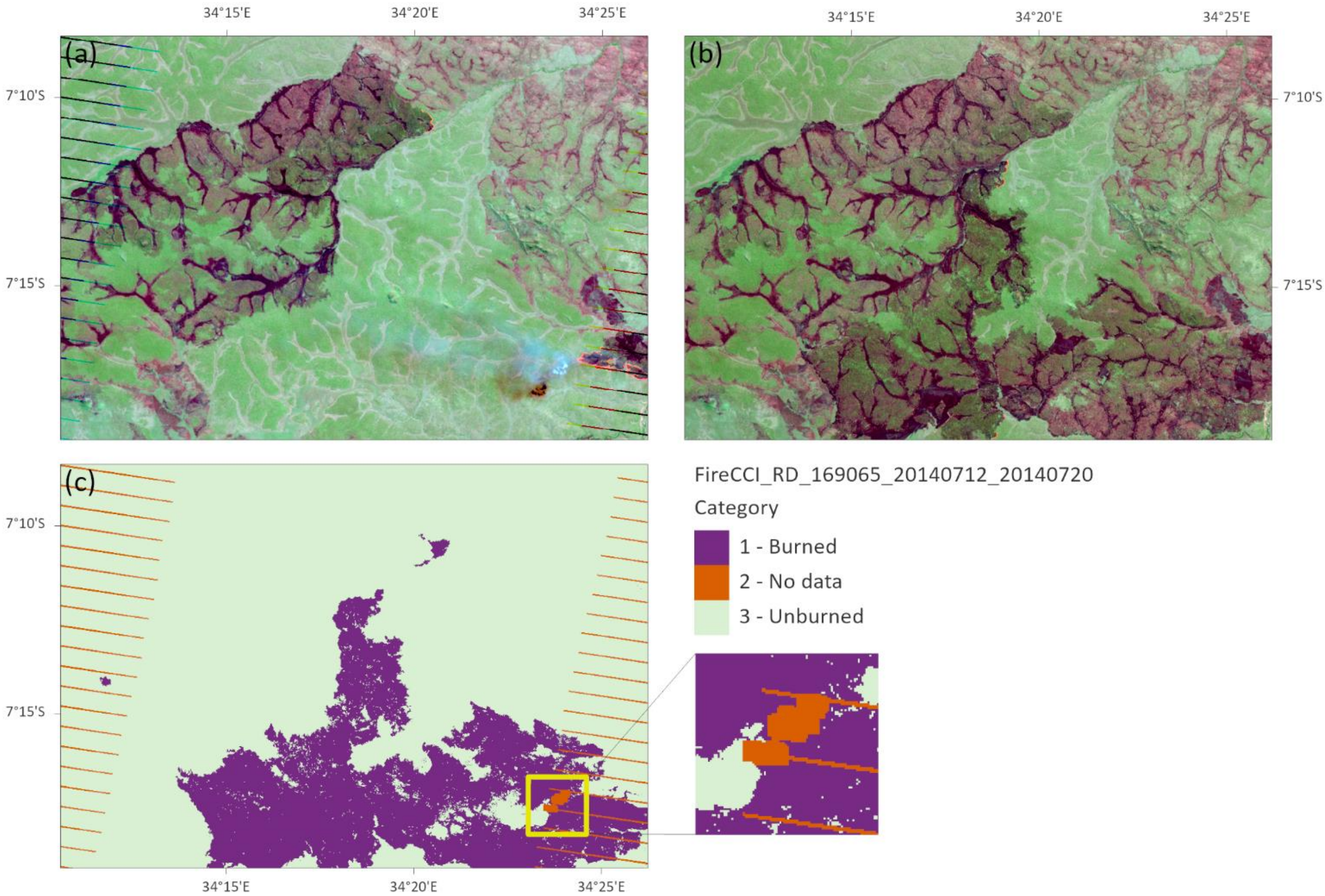


Figure 1. (a)Thiessen scene areas (TSAs) based on Landsat Worldwide Reference System-2 (WRS-2) frames. TSAs are used as non-overlapping spatial units in the sampling design. (b) Distribution of major Olson biomes reclassified as in Padilla et al. (2014).

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020



440 **Figure 2: Example of Landsat-7 pre-fire (a) RGB (7,4,3) image and Landsat-8 post-fire (b) RGB (7,5,4) image. Both (a, b), were used to derive the 'FireCCI_RD_169065_20140712_20140720' BA reference file (c) at WRS-2 Landsat 169-065 path-row (east Africa). Time distance between both images is 8 days: from 12 June to 20 June 2014. Only the land surface that burns between the two dates is classified as burned, while burned scars in the pre-fire image are assigned to the unburned category. Unobserved pixels on either pre- or post-image due to the presence of clouds, cloud-shadows, SLC-gaps or smoke plumes are classified as No data.**

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

Open Access Earth System Science Data Discussions

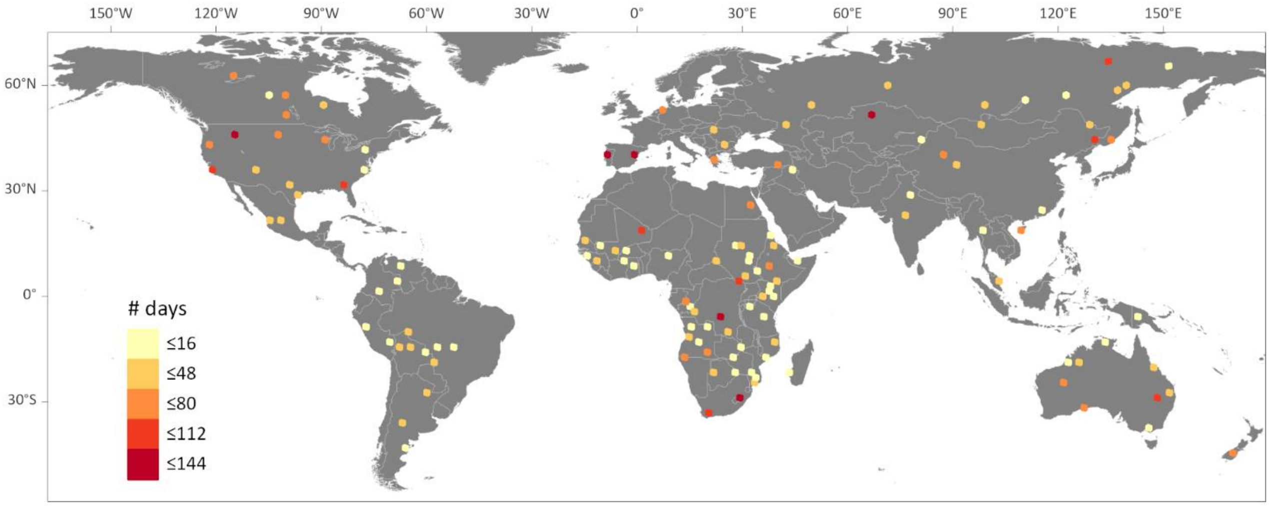
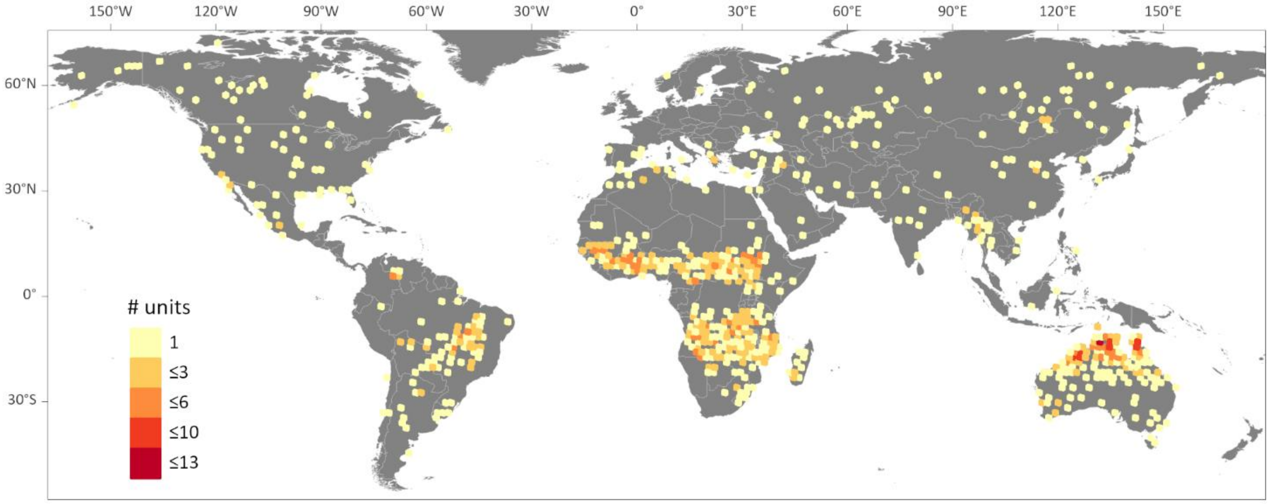


Figure 3: Spatial distribution of the reference sites for FireCCI global (2008) dataset. The legend shows the temporal distance (days) between the pre- and post-fire images used in each validation site for the year 2008.



445

Figure 4: Spatial distribution of the validation Thiessen scene areas (TSAs) for FireCCI global (2003-2014) dataset. The legend shows the total number of reference data files generated for each TSA between the period 2003-2014.

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

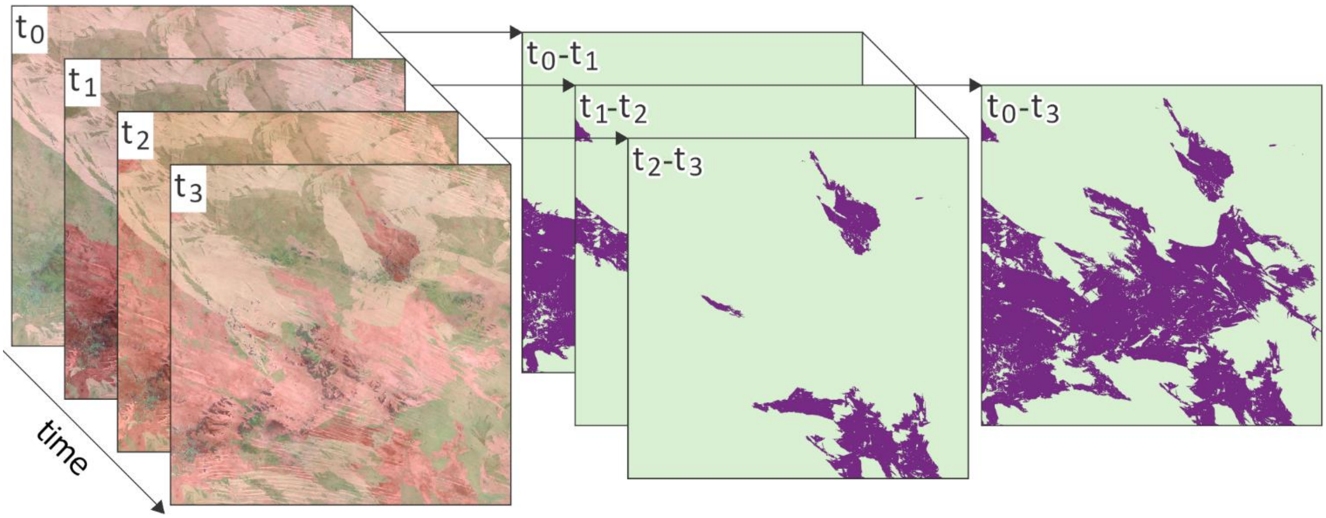


Figure 5: Schematic process of long units generation. Consecutive image pairs are selected from the multitemporal image series at same location (left: Landsat-8 RGB (7,5,4) images time series) to derive the correspondent short unit reference data files (e.g. Image t_0 and t_1 to obtain the reference data t_0-t_1). From the union of the different short units we generate the long unit reference data (right). The long unit t_0-t_3 includes all the burned scars occurred between the first image (t_0) and the last image interpreted (t_3).

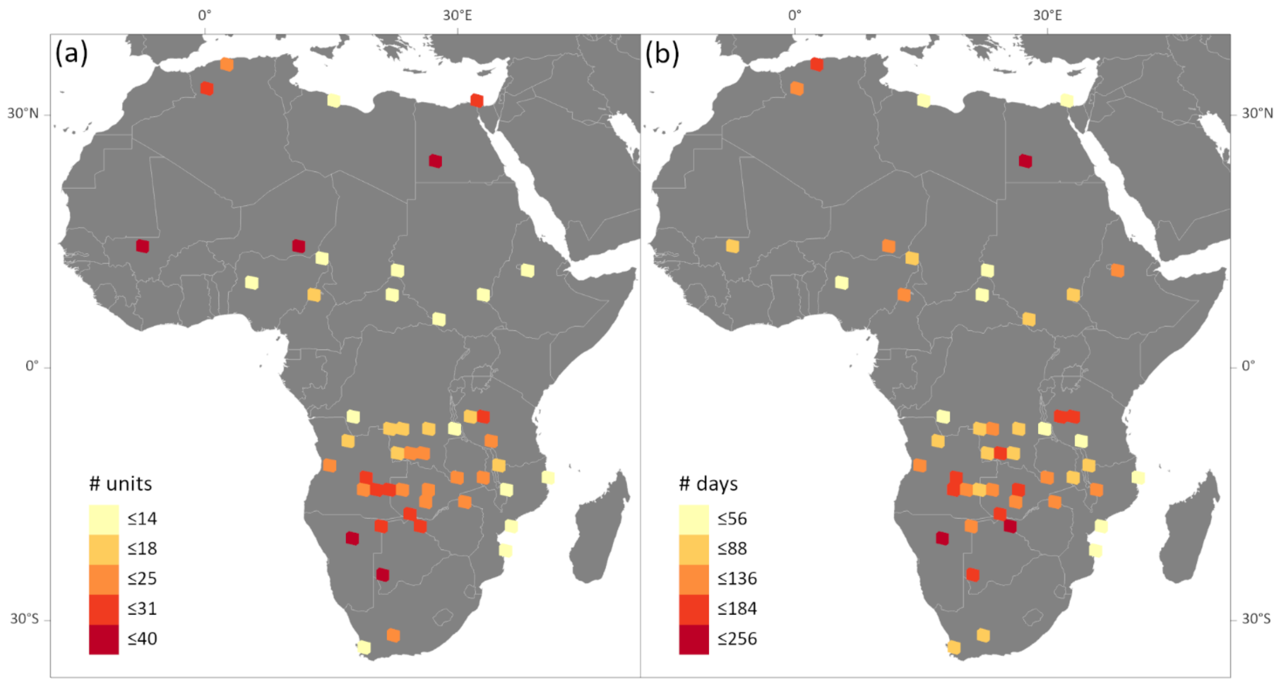


Figure 6: Spatial distribution of the reference sites for the FireCCI Africa (2016) dataset: (a) number of short units interpreted in each validation site and (b) temporal length of the long units.

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

Open Access Earth System Science Data Discussions

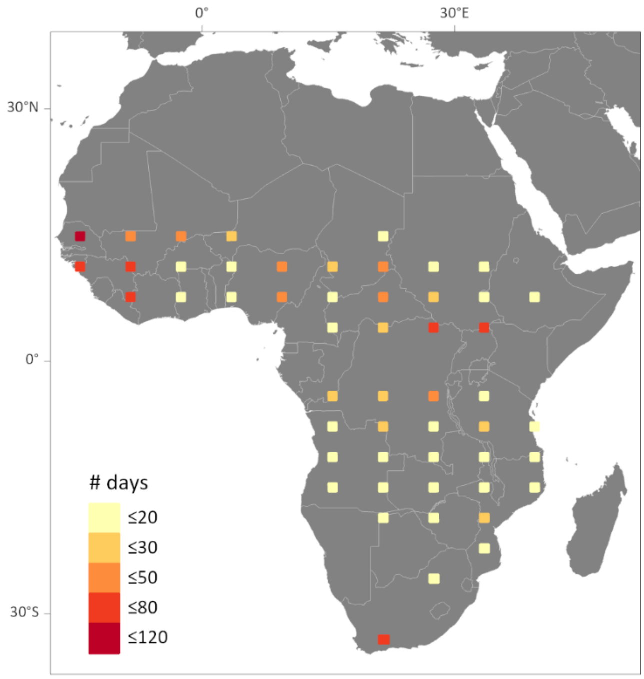


Figure 7: Spatial distribution of the reference sites for FireCCI Africa S2 (2016) dataset. The legend shows the temporal distance (days) between the pre- and post-fire images used in each validation site for the year 2016.

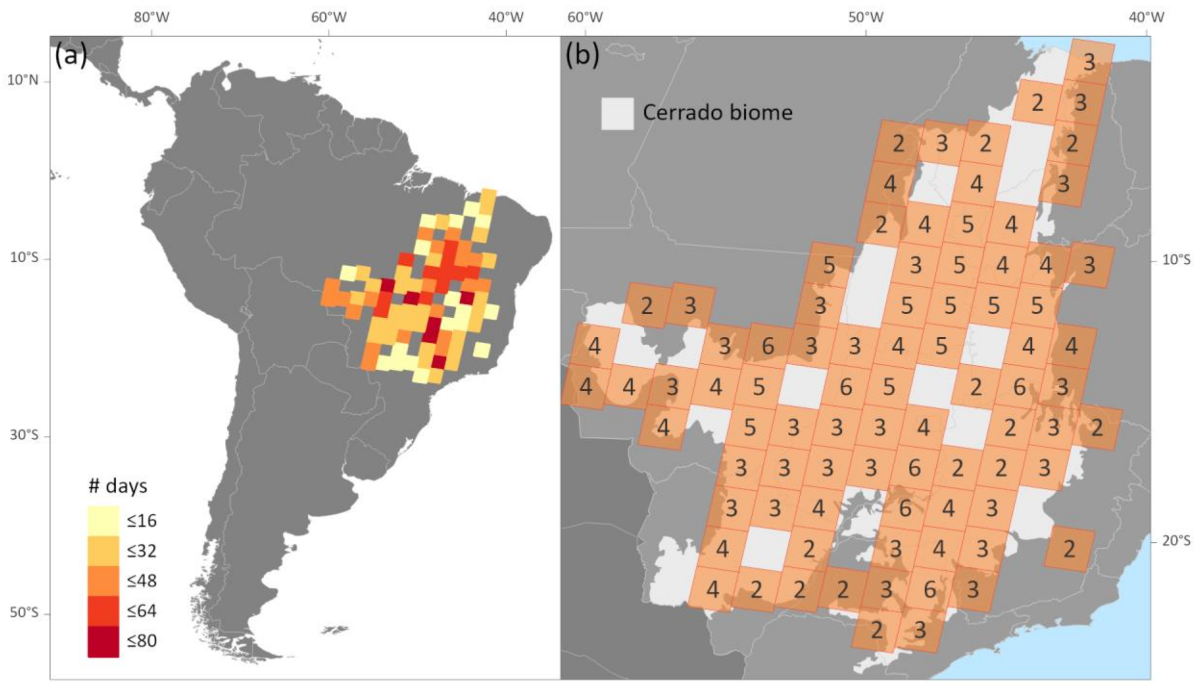


Figure 8: Spatial distribution of reference sites for BrFLAS Brazil (2015) dataset. (a) Temporal length of the reference files and (b) total number of images used in each validation site.

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

Open Access Earth System Science Data Discussions

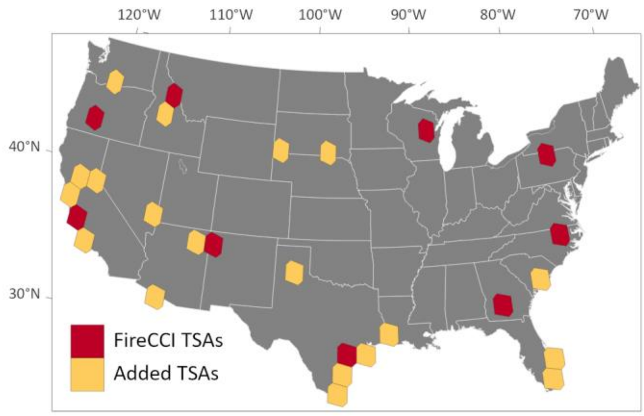


Figure 9: Spatial distribution of the validation Thiessen scene areas (TSAs) for BAECV CONUS (1988-2013) dataset. Modified from Vanderhoof et al. (2017).

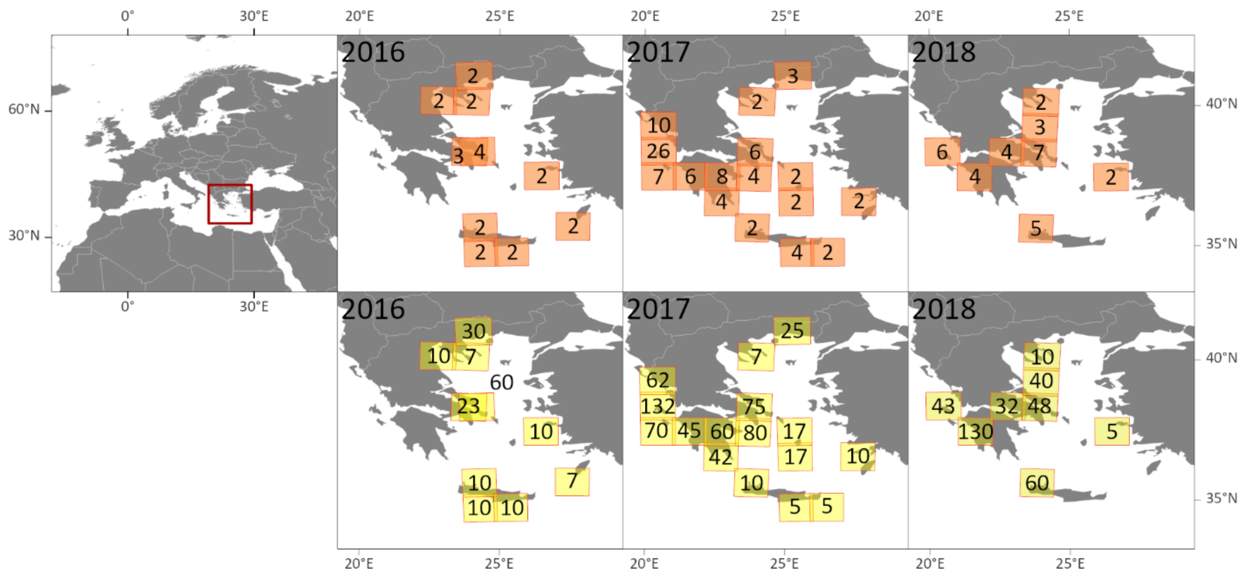


Figure 10: Spatial distribution of validation sites for NOFFi Greece (2016-2018) reference dataset based on Sentinel-2 tiles. The orange figures above show the number of images used in each validation site for each year, whereas the yellow ones below show the temporal length (days) of the reference data files generated in each validation site.

Table 1: Example of the standard attribute table of the reference shapefiles.

category	preDate	postDate	preImg	postImg	path	row	year	area
3	1988-07-05	1988-10-25	LT50150351988187XXX05	LT50150351988299XXX08	15	35	1988	267043.6
2	1988-07-05	1988-10-25	LT50150351988187XXX05	LT50150351988299XXX08	15	35	1988	4557.8
1	1988-07-05	1988-10-25	LT50150351988187XXX05	LT50150351988299XXX08	15	35	1988	2043.3
1	1988-07-05	1988-10-25	LT50150351988187XXX05	LT50150351988299XXX08	15	35	1988	900.4

https://doi.org/10.5194/essd-2020-74
Preprint. Discussion started: 27 April 2020

Open Access Earth System Science Data Discussions

Table 2: Datasets included in the Burned Area Reference Database. CCI: Climate Change Initiative, BrFLAS: Brazilian Fire-Land-Atmosphere System, BAECV: Burned Area Essential Climate Variable, NOFFi: National Observatory of Forest Fires, TM: Thematic Mapper, ETM+: Enhanced TM, OLI: Operational Land Imager, CEOS-LPVS: Committee on Earth Observing Satellites – Land Product Validation Subgroup.

Project	Reference Files (count)	Years	Extent	Source Imagery	CEOS-LPVS Stage
FireCCI global (2008)	129	2008	global	Landsat TM, ETM+	3
FireCCI global (2003-2014)	1200	2003-2014	global	Landsat TM, ETM+, OLI	3
FireCCI Africa	1102 (short and long units)	2016	Africa	Landsat ETM+, OLI	3
FireCCI Africa S2	52	2016	Africa	Sentinel-2	1
BrFLAS Brazil	84	2015	Brazil	Landsat OLI	1
BAECV CONUS	168	1988, 1993, 1998, 2003, 2008, 2013	United States	Landsat TM, ETM+, OLI	3
NOFFi Greece	34	2016-2018	Greece	Sentinel-2 MSI	1