eman ta zabal zazu

Universidad del País Vasco Euskal Herriko Unibertsitatea

GASTEIZKO INGENIARITZA ESKOLA
ESCUELA DE INGENIERÍA DE VITORIA-GASTEIZ

BAMT

Burned Area Mapping Tools in Google Earth Engine

Version 1.7 – User Guide

Release date: April 27, 2022

Ekhi Roteta*

Aitor Bastarrika

Department of Mining and Metallurgical Engineering and Materials Science
School of Engineering of Vitoria-Gasteiz
University of the Basque Country (UPV/EHU)

*   Correspondence: ekhi.roteta@ehu.eus

Index

1 Introduction

Aitor Bastarrika developed a software for burned area (BA) detection: Burned Area Mapping Software or BAMS (Bastarrika et al., 2014), formerly downloadable at https://bastarrika.wordpress.com/. It was a semi-automatic BA mapper based on ArcGIS, applying a two-phase algorithm on Landsat images. It required the user to download and save all the necessary Landsat images, and the process was dependent on the local server's processing capabilities. However, due to difficulties to maintain the software udpated, it is no longer available.

Google developed a cloud-computing platform to process satellite data: Google Earth Engine (GEE) (https://earthengine.google.com/, last accessed in April 2022) (Gorelick et al., 2017). This includes access to a large database of satellite datasets, no need to store big amounts of data and a powerful ability of data processing. Taking advantage of all this and following the idea of BAMS, several tools have been developed in Google Earth Engine. No input data need to be downloaded, the whole processing is carried out in the cloud, and only the final results are exported and optionally downloaded; the only requirement is a good internet connexion.

The original BAMS software generated reference data in a Landsat scene between two consecutive dates; here, we have expanded it to cover the entire process of BA mapping, while including also Sentinel-2 images to improve both the temporal and spatial resolutions. These tools, called **Burned Area Mapping Tools** or ***BAMT***, can be used by anyone with a Google account and signed up in the Google Earth Engine platform. The five main **BAMT** tools are (Figure 1):

1) **BA Cartography** tool: creates a BA map in the geographical extent and period defined by the user, via a supervised classification.
2) **VA** tool: selects validation sites based on a stratified random sampling methodology (Boschetti et al., 2016; Padilla et al., 2017, 2014), to assess a BA map.
3) **VA Dates** tool: after identifying the validation areas with the previous tool, this one just indicates which dates should be used in each validation area for creating reference perimeters.

4) **RP** tool: creates high-quality burned areas in a validation site between two single dates, via a supervised classification, to be used as reference data by the validation process.
5) **Assessment** tool: carries out the validation process, by comparing a BA map with reference data and by computing accuracy metrics.

There is also an extra tool, not related to BA mapping:

6) **Image Viewer**: to display Landsat and Sentinel-2 images.

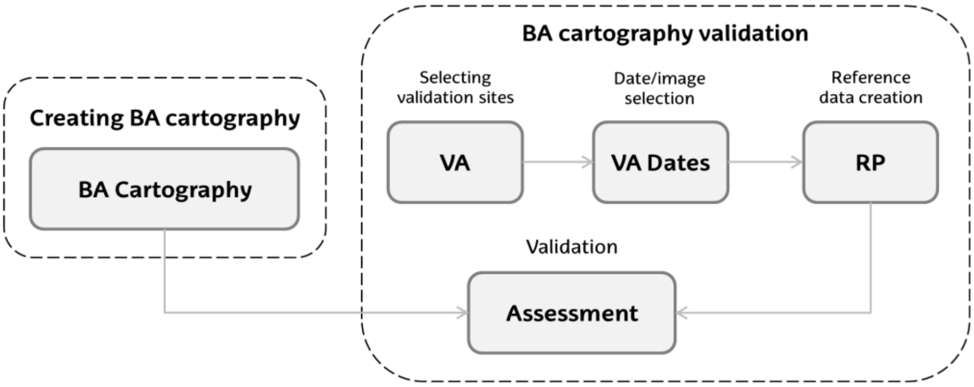


Figure 1. Flowchart of the five main BAMT tools.

All results are exported to the user's Google Drive account, where a folder called *BAMT_GEE* is created. Every file carries an identifier specified in each tool by the user, so that files from different projects can be distinguished.

The tools and the present user guide can be reached at https://github.com/ekhiroteta/BAMT. The study was presented as a scientific paper in the *Remote Sensing* journal, which should be consulted for further information on the tools' algorithms: https://www.mdpi.com/2072-4292/13/4/816. However, please keep in mind that this user guide corresponds to version 1.7, while 1.6 was the only available version when the paper was published.

Citation:

Roteta, E.; Bastarrika, A.; Franquesa, M.; Chuvieco, E. Landsat and Sentinel-2 Based Burned Area Mapping Tools in Google Earth Engine. *Remote Sens.* **2021**, *13*, 816. https://doi.org/10.3390/rs13040816

1.1 Datasets

Several datasets are used by these tools, all of them belonging to one of two large Earth observation satellite missions or programmes: Landsat and Sentinel.

The **Landsat** program is a NASA/USGS program for satellite imagery acquisition and Earth observation (https://www.usgs.gov/landsat-missions, last accessed in April 2022). From its eight historical and current satellites, the last five are used in these tools:

- **Landsat-4**, with the Thematic Mapper or **TM** sensor
- **Landsat-5**, with the Thematic Mapper or **TM** sensor
- **Landsat-7**, with the Enhanced Thematic Mapper Plus or **ETM+** sensor
- **Landsat-8**, with the Operational Land Imager or **OLI** sensor
- **Landsat-9**, with the Operational Land Imager 2 or **OLI-2** sensor

These five satellites currently cover a temporal span of 40 years, from July 1982 to the present, with each satellite acquiring images over the Earth at 30 m of spatial resolution every 16 days; the temporal resolution improves if different satellites are combined. Whenever one of these datasets is chosen, the Landsat Surface Reflectance (SR) product from Collection 2 will be used (https://www.usgs.gov/landsat-missions/landsat-collection-2-surface-reflectance, last accessed in April 2022), which consists in geometrically corrected images with Bottom of Atmosphere reflectances (BOA) for every band.

The **Sentinel-2** mission (S2) is part of the Copernicus Programme and developed by ESA (European Space Agency) for Earth Observation, (http://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2, last accessed in April 2022), consisting of twin satellites Sentinel-2A and Sentinel-2B. The same sensor, the Multi-Spectral Instrument or **MSI**, is carried by both satellites. The first satellite (Sentinel-2A) was launched in June 2015, so the mission currently covers a temporal span of 7 years, acquiring images first every 10 days and then every 5 days from March 2017 on (when the satellite Sentinel-2B was launched). The spatial resolution varies among 10, 20 and 60 m depending on the spectral band.

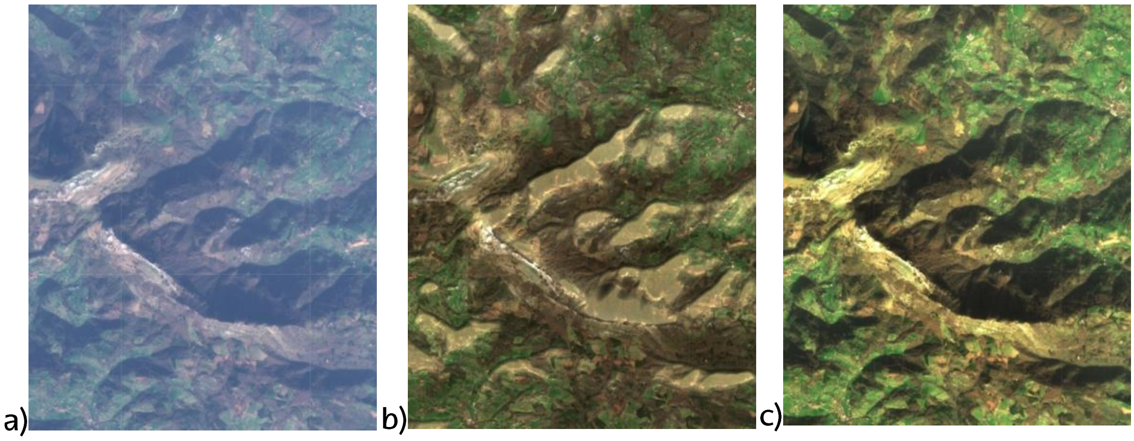


Figure 2. Sample area in a mountainous region from the 3rd of February 2020, tile 30TWN, where the artificial effect can be seen in mountain shadows. a) L1C product with TOA reflectances, b) L2A product with BOA reflectances as available in GEE, and c) the same L2A product generated independently with no topographic correction.

Whenever the Sentinel-2 dataset is selected in the BAMT tools, the Level-1C or L1C product is used (https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/product-types/level-1c, last accessed in April 2022), which contains orthorectified images with Top of Atmosphere (TOA) reflectances. The Level-2A or L2A product (https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/product-types/level-2a, last accessed in April 2022) contains Bottom of Atmosphere reflectances and an accurate classification of the image called Scene Classification (SCL), but it is not fully available yet in the Earth Engine Data Catalog. Besides, it was observed that some L2A images have some artificial effects in mountainous areas, resulting from a topographic overcorrection of mountain shadows (Figure 2), so the L1C product is preferred in these tools here.

Six spectral bands are used in all BAMT tools for both Landsat and Sentinel-2 datasets: the visible colours (blue, green and red), the near infrared (NIR) and two short wavelength infrareds (Short and Long SWIR). The near infrared B8A band, originally at 20 m, is used in most cases for S2 images; however, the B8 band at 10 m is selected as near infrared when the RP are created at that same resolution. Table 1 shows the corresponding bands in each satellite and sensor.

Table 1. Bands selected for BAMT tools, depending on the satellite and sensor.

Satellite	Landsat-4 and 5	Landsat-7	Landsat-8 and 9	Sentinel-2A and 2B	
Sensor	TM	ETM+	OLI and OLI-2	MSI	
Product	LSR c2			L1C	L2A
Blue	B1		B2	B2	
Green	B2		B3	B3	
Red	B3		B4	B4	
NIR	B4		B5	B8A (20 m) B8 (10 m)	
Short SWIR	B5		B6	B11	
Long SWIR	B7		B7	B12	
Quality band	QA_PIXEL			QA60	SCL

When clouds and cloud shadows need to be masked in any tool, quality bands are selected. For Landsat datasets the *QA_PIXEL* band is used, whose 3rd and 4th bits indicate the presence of cloud shadows and clouds, respectively. Sentinel-2 images have a similar quality band called *QA60*, with the same information in bits 10 and 11 (Table 1); an empirical threshold based on the B1 band is also employed, however, with every pixel B1 > 1500 being considered clouds.

However, the Scene Classification (SCL) is preferable for S2 images when the L2A is available, since the SCL has more categories and is spatially more accurate than the QA60 band. This is not feasible for the *BA Cartography* tool, where there may be thousands of images and coupling SCL bands from L2A products to the corresponding L1C images will exceed GEE's memory limit. But the SCL is used instead of the QA60 bands in the *RP* tool whenever the L2A product is available, where only two images are used in the whole process of the tool.

2 BA Cartography tool

https://code.earthengine.google.com/20c4db5b2e94290aff4b3a87cdb056c2

This main tool generates a burned area product over a large region and during a long temporal period, via a supervised classification. A polygon must be defined for the extent of the region, while the temporal period will be delimited by two dates.

The process consists of training a Random Forest classifier, using various spectral bands and indices from the post-fire image and from a multitemporal difference image. This classifier returns an image with a probability value from 0% (unburned) to 100% (burned) for every pixel. A two phase algorithm is applied upon this image, first identifying burned seeds (pixels with a very strong burned signal) and then extending the burned areas around these seeds (Bastarrika et al., 2011). The burned areas can be exported in the ESRI Shapefile format, with an attribute indicating the date of detection of the burned area.

2.1 Initial parameters

Before detecting burned areas, the user must define some initial parameters, which will fix the dates of the temporal periods, the geographic region and the dataset to be used. All these variables are defined at the beginning of the script, between lines 7 to 26.

The burned areas are detected comparing two temporal composites corresponding to the pre-fire and post-fire periods; the post-fire period begins exactly the day the pre-fire period ends. Thus, three different dates must be defined:

- ***date_1***: the beginning of the pre-fire period.
- ***date_2***: the end of the pre-fire period and the beginning of the post-fire period.
- ***date_3***: the end of the post-fire period.

All three dates must belong to the **String** data type and be written in this format:

YYYY-MM-DD

where YYYY stands for the year, MM for the month and DD for the day of the month, all of them differentiated with hyphens.

The dataset to be used for the product will be specified in the ***dataset*** variable. If '*Landsat*', all available images from sensors L4 TM, L5 TM, L7 ETM+, L8 OLI and L9 OLI-2 will be used at 30 m; otherwise, with the '*Sentinel*' value, data from the S2 MSI sensor, L1C product, will be chosen, with a 20 m spatial resolution.

The geographical region is defined manually as a polygon. There is a *Geometry Input* layer called ***studyArea*** for this purpose (Figure 3); if you are not familiar with the *Geometry tools*, please visit the page https://developers.google.com/earth-engine/guides/playground#geometry-tools (last accessed in April 2022). Any pre-existing polygon should be removed from this layer before defining a new region.

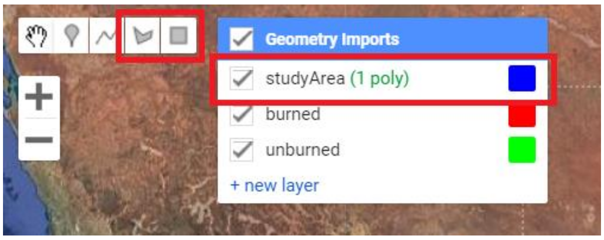


Figure 3. The *region_manual* layer to be used when the region must be defined manually as a polygon.

Finally, the user can write an ID in the ***identifier*** variable to differentiate this BA product from others, since this ID will appear in the names of the exported files. The last three variables, ***UMLError_tiles_1d***, ***UMLError_tiles_05d*** and ***UMLError_tiles_025d*** should be ignored and left as empty lists for now.

2.2 First layers

Once all the previous variables are defined, the user can press the button ***Run*** to execute the script.

The pre-fire and post-fire temporal composites are generated at this moment. For each period's composite, all available images from the region between the corresponding dates are selected, and clouds and cloud shadows are masked based on the quality band (*pixel_qa* in Landsat images, *QA60* for Sentinel-2 data). After computing the Normalized Burn Ratio or NBR spectral index, the composite is created by minimizing this index in every pixel. For the post-fire period another image is created, with the date from which each pixel was extracted. This image represents for every pixel the date when the lowest NBR value was registered, and will later be used as the date of detection of burned areas.

Six spectral bands are selected in both composites (Table 1): the three visible bands (blue, green and red), the near infrared (NIR), and both short wavelength infrareds (short and long SWIRs). Aside from those 6 bands, 3 spectral indices based on normalized differences are computed as well for both composites:

- Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1974)

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

- Normalized Burn Ratio (NBR) (Key and Benson, 1999)

$$NBR = \frac{NIR - LongSWIR}{NIR + LongSWIR}$$

- Normalized Burn Ratio 2 (NBR2) (García and Caselles, 1991)

$$NBR2 = \frac{ShortSWIR - LongSWIR}{ShortSWIR + LongSWIR}$$

At the same time, when the script is executed and the composites created, several layers are shown in the *Layer manager* and in the map:

- ***Exportable tiles***: the grid used later to export results, explained in section 2.4.
- ***Pre-fire***: a Long SWIR / NIR / Red colour composition of the pre-fire composite.
- ***Post-fire***: a Long SWIR / NIR / Red colour composition of the post-fire composite.
- ***Difference***: a simple difference between both colour compositions, $image_{post\text{-}fire}$ - $image_{pre\text{-}fire}$.
- ***MCD64A1*** and ***FireCCI51***: areas burned during the post-fire period according to the MCD64A1 c6 (Giglio et al., 2018, 2009) and FireCCI51 (Lizundia-Loiola et al., 2020) products, at 500 and 250 m respectively. As these products are only available from November 2000 and January 2001 on, the layers are not shown in the map for earlier dates.

A new button is printed in the *Console*:

- ***Zoom to region***: a general zoom to the region.

2.3 BA detection

Once some burned and unburned training polygons are defined using the available ***burned*** and ***unburned*** layers as *Geometry Imports* (Figure 4), burned areas can be detected for the first time. It is recommended to define polygons instead of rectangles, adjusting their shapes to burned areas; at the same time, defining large polygons can later slow down the process, so take care they are not too large. At least one training polygon is required for each class; otherwise a new polygon will be asked to be defined.

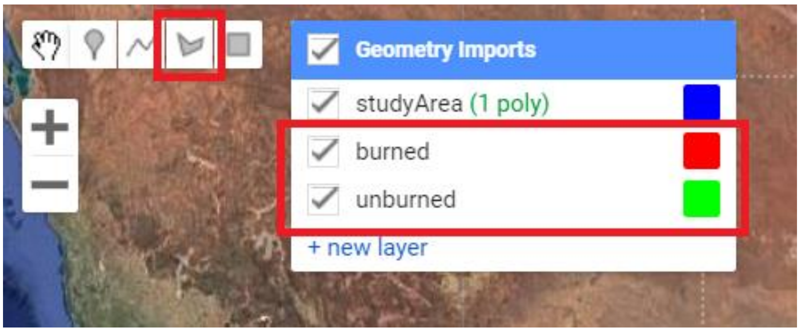


Figure 4. The layers to be used to define training polygons.

After pressing the ***Run*** button, burned areas are detected based on the training polygons. These areas are classified by a Random Forest classifier with 100 decision trees and nodes with a minimum of 10 leaves; the classifier is trained by all pixels contained in the training polygons. A total of 18 variables are introduced, the original 6 bands (blue, green, red, NIR and both SWIRs) and 3 spectral indices (NDVI, NBR and NBR2), each one of them twice: the post-fire value, and the difference between the pre- and post-fire values.

The result returned by the classifier is a probability image with values from 0% (unburned) to 100% (burned). Burned seeds are extracted from this image, representing pixels with the strongest burned signal; the threshold used here is the average of the mean burned probabilities in every burned training polygon. Two new layers are now added to the *Map* (Figure 5):

- ***BA***: the probability image. To make easier its visualisation, a threshold of 50% is applied to this image before being displayed, so that pixels below this threshold are transparent and the image at the bottom can be seen.

Burned pixels are displayed with a colour palette from white (probability values close to 50%) to blue (100% probability of burned).

- ***Seeds***: the seeds extracted from the previous probability image in a red colour.

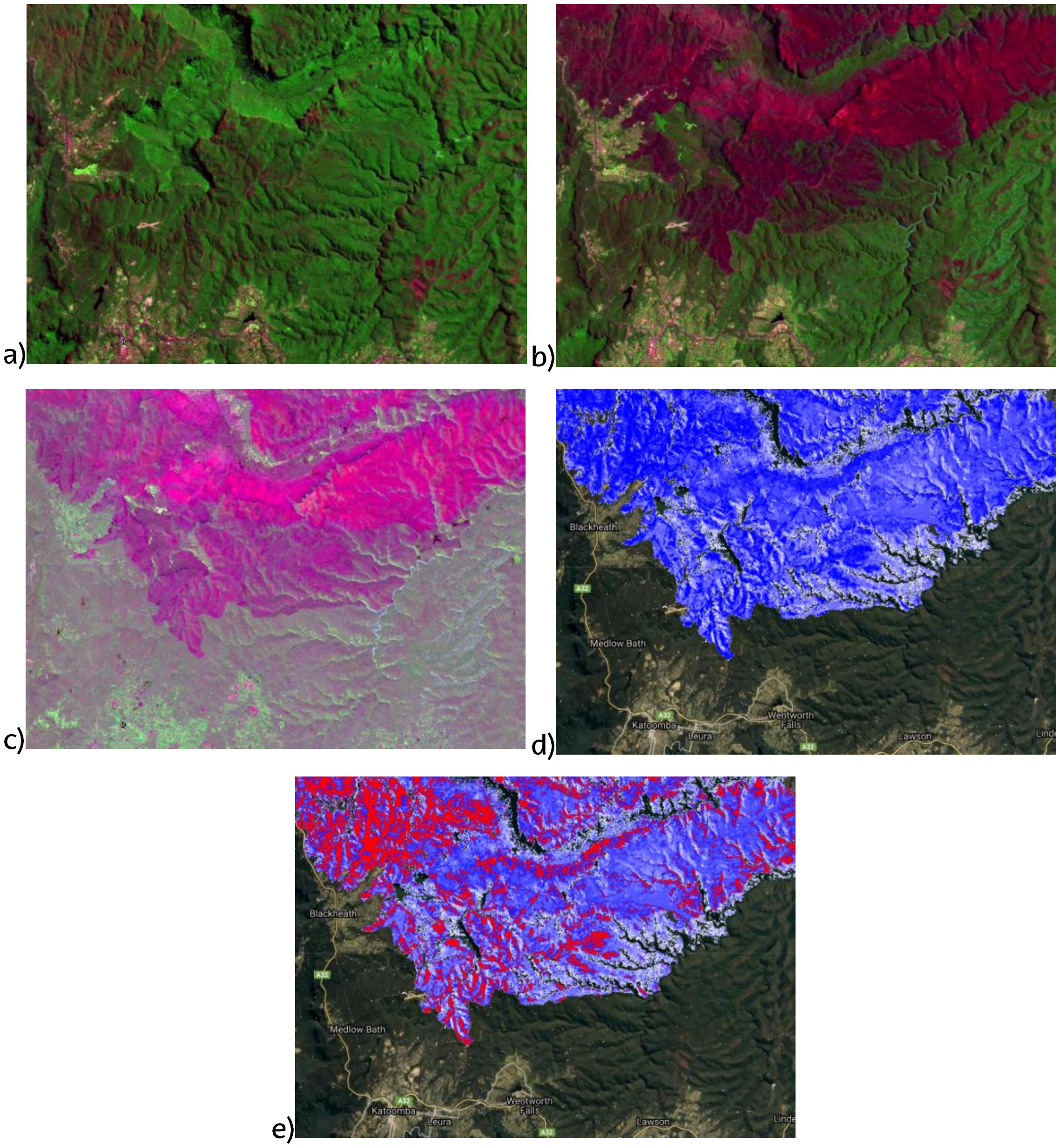


Figure 5. Sample area west of Sidney, in Australia, with Landsat data. a) pre-fire composite of the September-October 2019 period, b) post-fire composite of the November-December 2019 period, c) difference image between both periods, d) probability of burn image with the Google Satellite image in the background, and e) the seeds in red shown on the previous image.

The user can now modify or define new training polygons until the desired results are obtained, pressing the ***Run*** button at each iteration to update the BA. Not all pixels with a probability higher than 50% are exported in the end; burned areas are expanded from seeds outwards up to a 50% threshold (section 2.4). Thus, the user should keep in mind that burned patches in layer ***BA*** will not be exported if they do not contain any seed.

2.4 Exporting results

When the desired result is finally achieved, it can be exported to the user's Google Drive account by two new buttons in the *Console*:

- ***Export BA to Google Drive***: the whole final BA product is exported in ESRI Shapefiles. The polygons have only one attribute, *BurnDate*, which is the day of detection of the burned area in the YYYYMMDD format, where YYYY stands for the year, MM for the month and DD for the day of the month. Areas not observed in either the pre-fire or the post-fire period (usually corresponding to clouds and cloud shadows) are assigned a 0 value instead of a date; unburned pixels are not represented in these files.
- ***Export probability images***: the probability of burn image resulting from the Random Forest classifier is exported in GeoTIFF format, with values ranging from 0 to 100.

And a third button:

- ***Export BA as GEE assets***: the whole final BA product is exported as a vector layer, but as a GEE asset, instead of exporting it to the user's Google Drive account; they are stored in a folder called *BAMT_BA*. This option is recommended when the results are intended to be validated by the Assessment tool; otherwise, all ESRI Shapefiles exported to Google Drive will have to be uploaded to GEE instead.

Due to GEE's memory limits, the whole product cannot be exported in a single file when BA are detected in a large study area. To solve this, the product (both the shapefile and the probability image) is exported in a regular grid between latitudes 60°S and 80°N (Figure 6a). All tiles from this grid are 2 degrees high, but their width depends on the latitude: 2 degrees between latitudes 50°S-50°N, 3 degrees between 50°-70° in both hemispheres, and 6 degrees between 70°-80°N. There is an overlapping area of 0.1 degrees between adjacent tiles. The ID of each tile contains the latitude and longitude of the upper left corner (Figure 6b). This grid can be seen in the GEE map by activating the layer '***Exportable tiles***'.

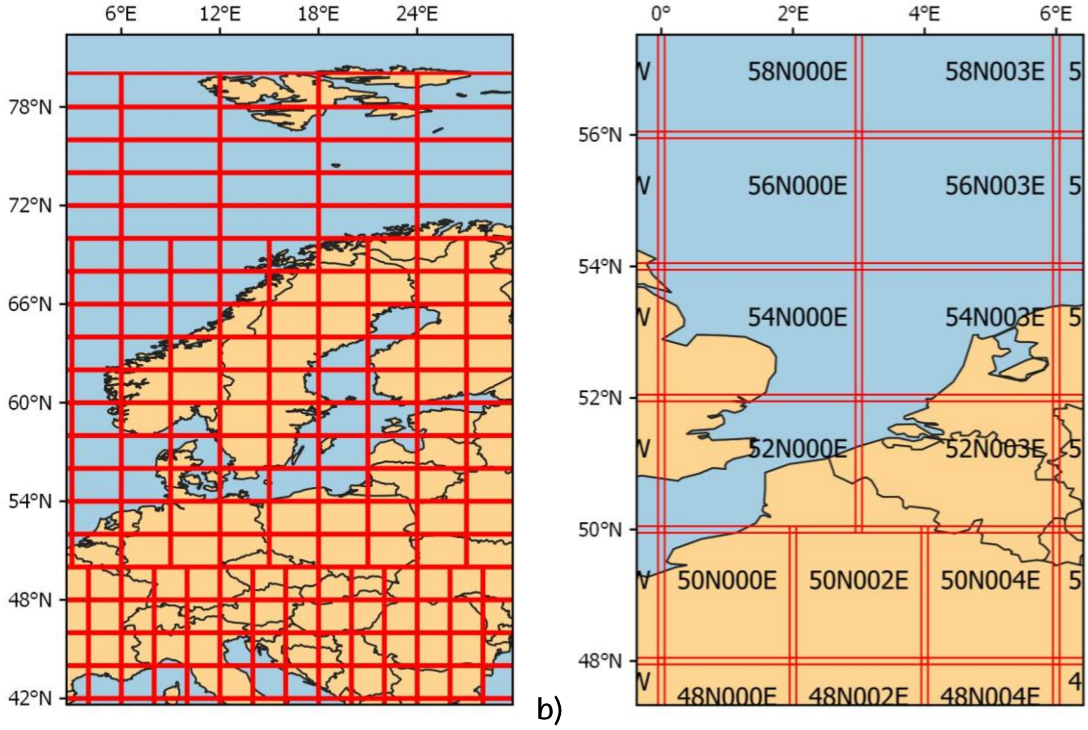
a) 

Figure 6. a) the composition of the grid in a sample area in Central and North Europe where different widths of tiles can be seen depending on the latitude, and b) a zoom in Western Europe where the tile IDs are shown.

Thus, these are the names of the exported files:

1) **BAMT_BA_*ID*_*SENSOR*_*DATE1*-*DATE2*_TILE-*TILEID*_SHP.shp**
2) **BAMT_BA_*ID*_*SENSOR*_*DATE1*-*DATE2*_TILE-*TILEID*_TIF.tif**

where:

- **ID**: the identifier defined by the user.
- **SENSOR**: *S2MSI* or *Lndst*, depending on the used dataset, corresponding to Sentinel-2 MSI and Landsat respectively.
- **DATE1**: the beginning of the post-fire period, in *YYYYMMDD* format.
- **DATE2**: the end of the post-fire period, in *YYYYMMDD* format.
- **TILEID**: the ID of the exported tile, indicating the latitude and longitude of the upper left corner.

After pressing the button ***Export BA*** or ***Export probability images***, the files to be exported will appear in the *Tasks* tab. The user must press the button ***Run*** next to these files or tasks one by one, and they will be exported to the user's Google Drive account, in a folder called *BAMT_GEE*, or to the asset collection in GEE, in a folder called *BAMT_BA*.

As it was told before, there may be some slight differences between the burned areas displayed in the GEE API map and those exported in shapefiles to Google Drive or as GEE assets. A two-phase algorithm is applied on the RF probability image before vectorizing it, with the aim to reduce noise. In the first phase burned seeds are extracted from the probability image, as already explained. In a second phase, the burned areas are expanded from these seeds outwards up to a 50% threshold. These grown burned areas are then polygonized and exported.

The post-fire temporal composite was created by minimizing the NBR spectral index over the whole period, which means that every pixel may have its origin in a different date. As the final burned polygon should only have one day of burn, the modal date among all the pixels is assigned to the polygon, i.e. the most repeated date. This is the value represented by the *BurnDate* attribute. Polygons split between two adjacent tiles may display different dates of burn, even if they are clearly part of the same burned area, because they are composed by different pixels and the most repeated date may have changed.

Despite splitting the BA product in 2 degree tiles, the burned polygons in a single tile are sometimes too heavy to export, and the processing fails with the task turning red and showing a message of '*Error: User memory limit exceeded*' or '*Error: Computation timed out*' (Figure 7).

a)

b) 

Figure 7. Errors shown every time '*user memory limit exceeded*' or '*computation timed out*'.

In those cases the 2 degree tile must be split in smaller subtiles; this is when the ***UMLError_tiles_1d*** variable must be used. The IDs of the tiles that could not be exported should be written in this list, as *String* variables, and after running the script these tiles will be split in 4 smaller tiles, 1 degree high and wide (Figure 8). The new subtiles will be named with an index from 1 to 4 hyphenated to the ID of the parent tile.

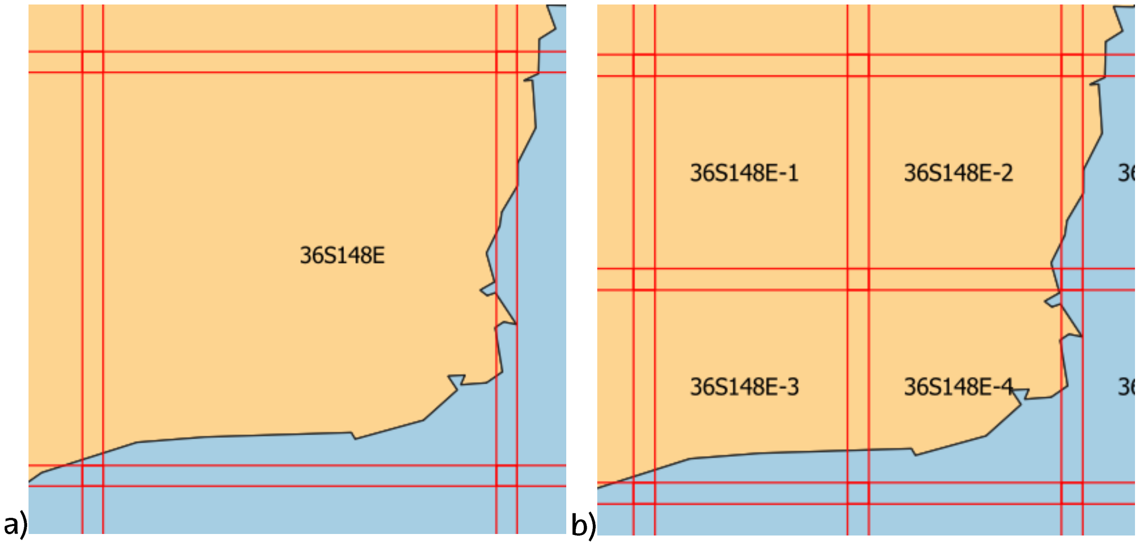


Figure 8. a) the original 2 degree tile with the upper left corner at latitude 36°S and longitude 148°E, in Southeast Australia, and b) the same tile split in four smaller subtiles.

If some of these 1 degree tiles cannot be exported yet, they can be split again in 0.5 degree tiles writing them in the ***UMLError_tiles_05d*** variable; or even farther by writing the undownloaded 0.5-degree tile in ***UMLError_tiles_025d*** and thus getting subtiles 0.25 degrees high and wide. The resulting grid can be seen at any moment in the map as the '***Exportable tiles***' layer (Figure 9).

a)
```
// Please list here the tiles that could not be downloaded because 'User memory limit
// exceeded':
var UMLError_tiles_1d = ['34S148E'];
// In case some 1 degree tiles cannot be downloaded yet, list them here:
var UMLError_tiles_05d = ['34S148E-3'];
// And if some 0.5 degree tiles cannot be downloaded yet, list them here:
var UMLError_tiles_025d = ['34S148E-3-3'];
```

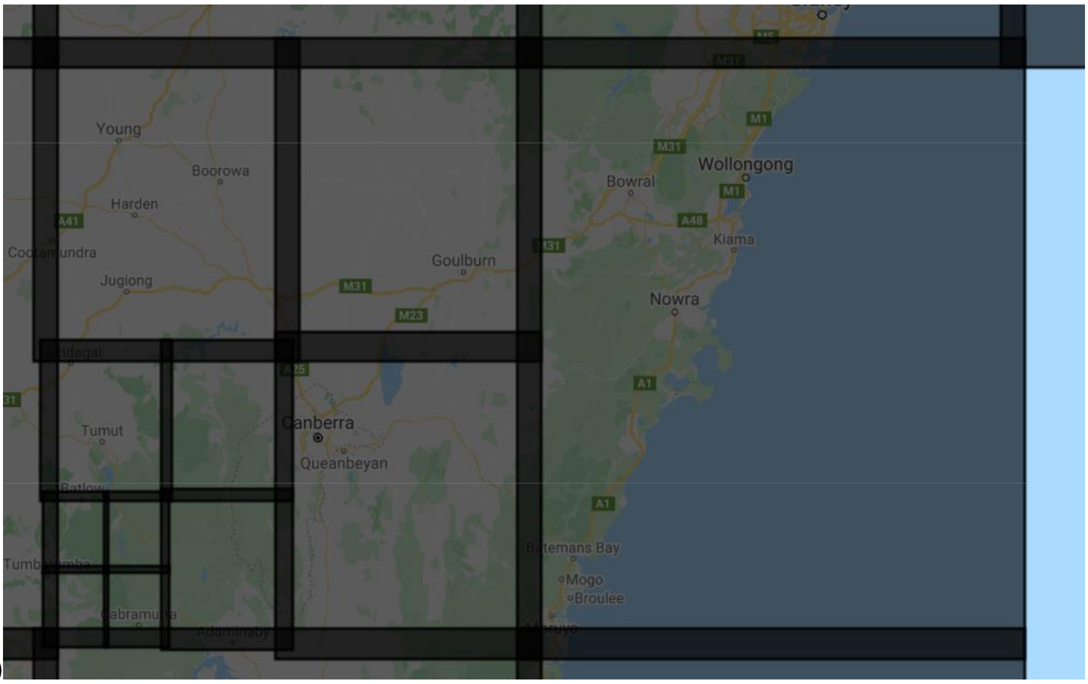
b)


Figure 9. a) Example of splitting a tile several times in the script, up to the point where subtiles 0.25 degrees high and wide are gotten; the original 2 degree tile's upper left corner is at 34°S 148°E around Canberra, Australia. b) Result of this splitting seen in the map; the tile to the East (ID 34S150E) remains unchanged.

Another common error is shown when every pixel in the entire tile is observed and unburned, so there is no polygon to be exported, which would be either burned or unobserved areas. The task turns red with a message telling '*Error: Unable to write shapefile from empty collection*' (Figure 10). This shapefile is simply not exported to the user's Google Drive account or GEE asset collection.

Finally, once the BA detected and all the necessary data exported, it is important not to forget to delete some polygons before beginning to detect BA elsewhere. The *studyArea* layer should not have more than a polygon at the same time, so the previous polygon should be removed before changing the region. As for burned and unburned training polygons, the tool filters these polygons spatially, so it is not necessary to delete them if the region is changed. They must be deleted, however, if the dates are changed but not the region, as pixels in training polygons will hardly correspond to the same burned/unburned class.

However, the training polygons can be saved in case the process must be repeated. It is as simple as saving a copy of the script with the *Save as…* button. While the user continues working in the original script, the region, the dates and the training polygons will be kept safe in the copy.

Figure 10. Error shown when there is no polygon to be exported.

3 VA tool

https://code.earthengine.google.com/d39add8c5de8904433500bd738f3dab6

Reference data need to be created for a BA map's assessment in some validation areas (VA). This tool selects the best validation areas to assess the product, adapting a stratified random sampling methodology (Boschetti et al., 2016; Padilla et al., 2017, 2015) in Google Earth Engine. Later, the *RP* tool will be used to create reference data in these validation areas, derived either from Landsat or Sentinel-2 data.

3.1 Sampling units

The original stratified random sampling methodology used Thiessen scene areas (TSA) (Cohen et al., 2010; Kennedy et al., 2010) based on Landsat scenes as sampling units. These cannot be used however with Sentinel-2 images, as many Landsat TSAs may be divided among several Sentinel-2 tiles or located on the edge of a S2 orbit swath, so S2 tiles are used instead when the user selects S2 data for creating reference perimeters.

Besides, both the TSAs and the S2 tiles are too large to process when creating reference data (around 170x150 and 110x110 km^2, respectively), so generating reference perimeters in a smaller central area of the TSA/tile is more suitable, as it has been done in other studies (Chuvieco et al., 2018; Franquesa et al., 2020; Lizundia-Loiola et al., 2020; Roteta et al., 2019); it will also be easier to process for GEE. Thus, squares of 100x100, 50x50, 30x30, 20x20 and 10x10 km^2 have been created in the centre of every TSA and S2 tile to be used as sampling units in BAMT (Figure 11); this way the user will choose the size of the validation areas. Note that the largest units, of 100x100 km^2, are nearly the same size as the original S2 tiles (110x110 km^2).

To avoid S2 sampling units divided between two orbits, units were removed when no orbit swath covered at least 95% of the square's surface. Units far from the Equator may be covered by two or more orbits, increasing the temporal resolution.

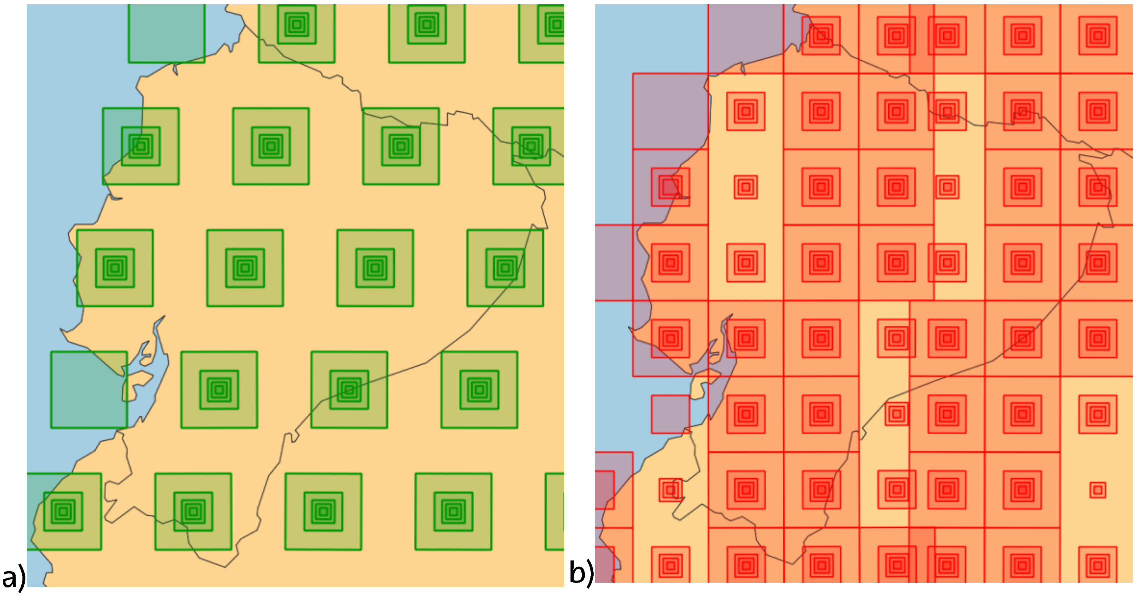


Figure 11. Sampling units for a) Landsat and b) Sentinel-2 datasets over Ecuador. Gaps between S2 sampling units are caused by orbit swath edges.

3.2 Description of the tool

As the validation areas are needed to assess the BA product, the same temporal and spatial extensions as those of the product should be defined here. The ***date_pre*** and ***date_post*** variables should coincide respectively with ***date_2*** and ***date_3*** of the *BA* Cartography tool (if the product was made by that tool), all of them in the '*YYYY-MM-DD*' format. In the same way, the poygon delimiting the BA product in the layer ***studyArea*** should be the same. Any pre-existing polygon should be removed from this layer before defining a new region.

The ***dataset*** variable will specify which sensor's data are to be used in the sampling of validation areas (and later when creating reference perimeters): Landsat-4 TM, Landsat-5 TM, Landsat-7 ETM+, Landsat-8 OLI, Landsat-9 OLI-2, all the previous five sensors (Landsat-4 to 9), or Sentinel-2 MSI. The user can write an ID in the ***identifier*** variable to differentiate this BA product from others, just as in the previous tool. The number of VA to sample is set by ***numberVA***, and the size (width and height) in km of each sampling unit and validation area by ***dimension***.

The original stratified random sampling methodology applies two stratification criteria:

1) The predominant Olson biome (Olson et al., 2001). The original 14 biomes are classified into seven categories to simplify the process.
2) The fire activity in each sampling unit. In this case either the MCD64A1 c6 (Giglio et al., 2018, 2009) or the FireCCI51 (Lizundia-Loiola et al., 2020) product is used as a reference, which is specified by the user with the ***globalBA*** parameter.

However, several stratification criteria for data availability can be used in the tool, first introduced by Boschetti et al., 2016. These criteria are optional, applied only when the variable ***availableData*** is **true**; if **false**, only the first two criteria (Olson biomes and fire activity) are used. Including these optional criteria may slow down the sampling process or exceed the memory limit with large study areas or long temporal periods, but it should be included as far as possible.

With these criteria, the user can require the sampling units' temporal periods to have a minimum length of ***minLength*** days, during which there is an available image every ***minFreq*** days, each one of them with a cloud coverage lower than the given ***maxCloud*** threshold. This way a continuous and frequent observation of the ground is ensured.

Unfortunately, computing the cloud coverage in every sampling unit and every single image in the temporal period requires too much processment for the GEE, so this information is taken from the image's metadata, which refers to the whole image. This means that the value used here is not exactly that of the sampling unit's area; but it is a good approximation.

Once all these parameters have been set properly, the script can be ***Run***, and three buttons are printed in the *Console*:

1) ***Zoom to region***: a general zoom to the region.
2) ***Show results***: processes the sampling methodology and shows the results.
3) ***Export results***: processes the sampling methodology and exports the results.

Upon pressing the button ***Show results***, the sampling methodology is carried out. If the optional criteria for data availability are applied, sampling units fulfilling the three conditions (period's minimum length, minimum frequency and maximum cloudiness) will be filtered, removing those that do not; the remaining units are assigned the temporal period where the conditions were achieved. There may happen to be two or more sampling units in a single Landsat scene or S2 tile, each with its own period, if the BA product's period is long enough. If ***availableData*** is **false**, all the sampling units are kept, and the BA product's period is assigned to them (from ***date_pre*** to ***date_post***).

The other two stratification criteria are applied on the sampling units resulting from the optional criteria. Each unit is assigned 1) the predominant Olson biome covering its area, and 2) the total burned surface according to the MCD64A1 or FireCCI51 product in the sampling unit's area during its period. Sampling units are divided between low and high BA substrata in each biome, and validation areas are selected from each BA/biome substratum in proportion to the number of units and total burned surface in that substratum.

When the processing is done, the list of the biomes found in the study area is printed in the *Console* under the buttons, as well as the sampling units and the sampled validation areas. And three layers are added to the *Map*:

1) ***MCD64A1*** or ***FireCCI51***: the burned areas according to the selected product between ***date_pre*** and ***date_post***.
2) ***Sampling units***: the sampling units, filtered by the optional criteria for data availability if these were applied, before the stratified random sampling methodology is run.
3) ***Validation Areas***: the resulting validation areas sampled in this process.

The tool may have sampled fewer validation areas than specified by the ***numberVA*** variable, due to an insufficient amount of sampling units in a BA/biome stratum. In this case the user should increase the number of VA to sample, and re-run the process.

To export the results to the user's Google Drive account the button ***Export results*** must be used.

If the study area is too large or the period of the BA product too long, the '*User memory limit*' may exceed after pressing the ***Show results*** button. In this case it is more preferable to use directly the button ***Export results***, without displaying the results in the *Map*, since GEE grants more memory capacity to processes that are exported but not shown in the interface. Not applying the optional criteria also helps making the process faster.

Two files can be exported, which have the following names:

1) **BAMT_VA_*ID_SENSOR_DATE1-DATE2_DIM*km_SU.shp**
2) **BAMT_VA_*ID_SENSOR_DATE1-DATE2_DIM*km_VA.shp**

where:

- **ID**: the identifier defined by the user.
- **SENSOR**: *S2MSI* or *Lndst*, depending on the selected dataset, corresponding to Sentinel-2 MSI and Landsat respectively.
- **DATE1**: the beginning of the BA product's period (***date_pre***), in *YYYYMMDD* format.
- **DATE2**: the end of the BA product's period (***date_post***), in *YYYYMMDD* format.
- **DIM**: the width and height of the validation areas, which are squares in the centres of Landsat scenes or S2 tiles.

These two ESRI shapefiles correspond to the layers containing sampling units (SU) and sampled validation areas (VA), respectively, as already displayed in the GEE interface. Every feature in the shapefile has these attributes below:

- ***BIOME***: the predominant Olson biome found in the area.
- ***DATE1***: the beginning of the sampling unit's period, in *YYYYMMDD* format.
- ***DATE2***: the end of the sampling unit's period, in *YYYYMMDD* format.
- ***DAYS***: the length of the period in days, from ***DATE1*** to ***DATE2***.
- ***BA***: the total burned surface in km^2 in the unit between ***DATE1*** and ***DATE2*** according to the MCD64A1 or FireCCI51 product.
- ***BAstratum***: the stratum this unit belongs to (high and low BA strata).

There will be two more attributes, which depend on the chosen dataset. If Landsat was chosen:

- ***PATH***: the path of the Landsat scene the sampling unit was based on.
- ***ROW***: the row of the Landsat scene the sampling unit was based on.

Or if Sentinel-2 was chosen:

- ***TILE***: the code of the S2 tile the sampling unit was based on.
- ***ORBIT***: the codes of orbits whose swaths cover at least 95% of the validation area.

4 VA Dates tool

https://code.earthengine.google.com/95aaacf6b9e2c29f922c421472724c24

After sampling the validation areas, the next step is to create reference data, but the exact dates where images are available must be known for that. This tool simply tells and shows the user the available dates in each validation area.

The tool works with an interface, which makes it easier to handle; this interface will appear after pressing the button ***Run***. The parameters that must be defined are divided in four sections:

1) ***Dataset***: the data must be chosen from a list that contains all the possible values (the same as in the previous tools).
2) ***Region***: the size of the validation area (10, 20, 30, 50 or 100 km), and the path and row (if a Landsat dataset was chosen) or the tile code (if the Sentinel-2 dataset was chosen).
3) ***Dates***: the beginning and end of the validation area's period.
4) ***Cloud coverage***: images with a cloud coverage higher than this value will not be displayed.

The dataset, the size of the validation area and the cloud coverage should be the same as in the previous VA tool. The path/row or tile and the dates will depend on the validation area.

After defining all these parameters, the button ***Load and display*** can be pressed, and all the available images in the validation area for the defined period will be added in the *Map*, all of them with a cloud coverage lower than the defined threshold (according to the metadata); the date of each image can be seen in the *Layers* list in the *YYYYMMDD* format (Figure 12). Please keep in mind that the cloud coverage value in the image's metadata is used, so images from some dates may contain too many clouds despite the filtering threshold.

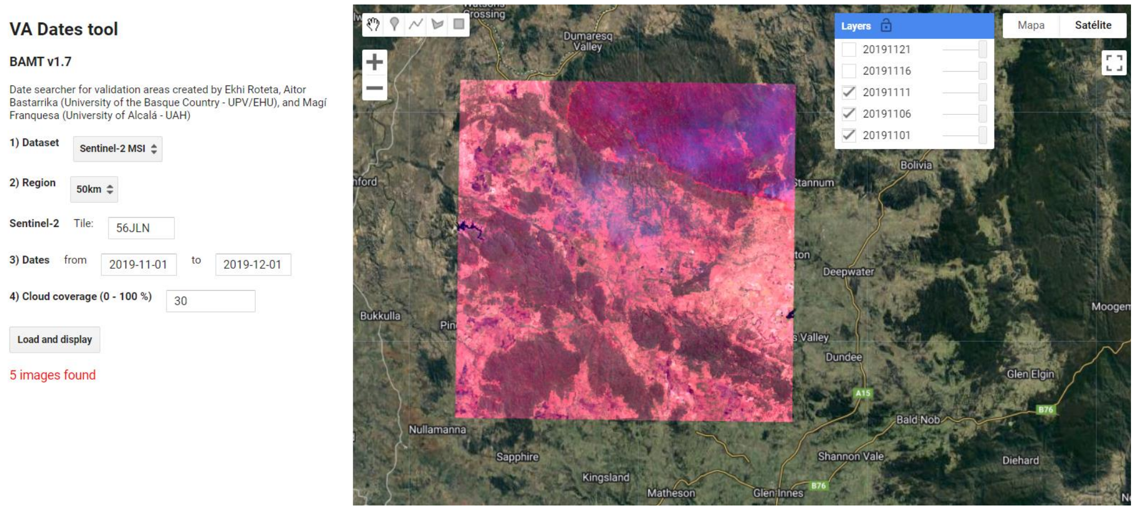


Figure 12. Example image of the *VA Dates* tool; input parameters can be seen in the interface on the left, and the resulting images in the map.

5 RP tool

https://code.earthengine.google.com/5864c922d6171bcb85666ddb4451f6fd

The aim of this tool is a high-quality burned area detection between two different images for assessment purposes, via a supervised classification; the images are selected from pre-fire and post-fire single dates defined by the user, instead of using temporal composites.

The process, practically identical to that of the *BA* Cartography tool, consists in training a Random Forest classifier, using various spectral bands and indices from the post-fire image and from a multitemporal difference image. This classifier returns an image with a probability value from 0% (unburned) to 100% (burned) for every pixel. A two phase algorithm is applied upon this image, first identifying burned seeds (pixels with a very strong burned signal) and then extending the burned areas around these seeds. The burned areas can be exported in the ESRI Shapefile format, with only two possible values for the unique attribute: *BURNED* or *UNOBSERVED*.

5.1 Initial parameters

In the first step of the burned area detection, the user must define some initial parameters, which will fix the dates, the geographic region and the dataset to be used. All these variables are defined at the beginning of the script, between lines 7 to 37.

5.1.1 Dates

The burned areas are detected comparing two different images corresponding to the pre-fire and the post-fire dates:

- ***date_1***: the pre-fire date.
- ***date_2***: the post-fire date.

All three dates must belong to the **String** data type and be written in the 'YYYY-MM-DD' format. The user can use the *VA Dates* tool to know the available dates.

5.1.2 Region

There are three different options for the ***region_type*** variable: '*Landsat*', '*Sentinel*' or '*manual*'.

Landsat

The region is a square in the centre of a Landsat scene. The scene is selected by the ***path*** and ***row*** variables, both in a **Number** (and not String) data type; the dimension of the region is defined by the ***region_dimension*** variable, which is the width and height of the square in km; only values of *10*, *20*, *30*, *50* or *100* km are allowed for this dimension. Only Landsat data are used in this option, from Landsat-4 on (Landsat-4 TM, Landsat-5 TM, Landsat-7 ETM+, Landsat-8 OLI and Landsat-9 OLI-2). As the Landsat-4 satellite was launched in July 1982, there is no available image before that date. Burned areas are processed at 30 m spatial resolution.

Sentinel

Similar to the *Landsat* region type, this region is a square in the centre of a Sentinel-2 tile, selected by the ***tile*** variable in a **String** data type. The ***region_dimension*** variable defines the width and height of the square in km, with allowed values of *10*, *20*, *30*, *50* and *100*. Only Sentinel-2 MSI data are used in this option, combining both Sentinel-2A and 2B satellites when available; as the first satellite (S2A) was launched in June 2015, there is no available image before that date. Burned areas can be processed at both 20 and 10 m spatial resolution, defined by the ***resolution*** variable.

manual

This option is recommended when the user needs to detect burned areas in an area not located in the centre of a Landsat scene or a Sentinel-2 tile. In this case the region is defined manually as a rectangle; there is a *Geometry Input* layer called ***region_manual*** for this purpose. Any preexisting polygon should be removed from this layer before defining a new region.

When this manual region is selected, the user must specify which dataset will be used, Landsat or Sentinel-2. If Landsat data are selected, the burned areas will later be detected at 30 m; otherwise, with Sentinel-2 data, at 20 or 10 m, depending on the defined value for the ***resolution*** variable.

5.2 First layers

Once all the previous variables are defined, the user can press the button ***Run*** to execute the script.

At this moment, the pre-fire and post-fire images are taken from the dates defined by the user, selecting six spectral bands from each image (Table 1): the three visible bands (blue, green and red), the near infrared (NIR), and both short wavelength infrareds (short and long SWIRs). Aside from those 6 bands, 3 spectral indices based on normalized differences are computed for both images as well:

- Normalized Difference Vegetation Index (NDVI)

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

- Normalized Burn Ratio (NBR)

$$NBR = \frac{NIR - LongSWIR}{NIR + LongSWIR}$$

- Normalized Burn Ratio 2 (NBR2)

$$NBR2 = \frac{ShortSWIR - LongSWIR}{ShortSWIR + LongSWIR}$$

At the same time, several layers are shown in the map, which can be managed in the *Layer manager*:

- ***Pre-fire***: a colour composition of the pre-fire image, with the Long SWIR, NIR and Red bands corresponding to the RGB colours.
- ***Post-fire***: a colour composition of the post-fire image, with the Long SWIR, NIR and Red bands corresponding to the RGB colours.
- ***Difference***: a simple difference between both colour compositions, $image_{post\text{-}fire} - image_{pre\text{-}fire}$.
- ***MCD64A1***: areas burned between the two dates according to the MCD64A1 c6 product, at 500 m. As this product is only available from November 2000 on, this layer is not shown in the map for earlier dates.
- ***MCD64A1*** and ***FireCCI51***: areas burned between the pre-fire and post-fire dates according to the MCD64A1 c6 (Giglio et al., 2018, 2009) and FireCCI51

(Lizundia-Loiola et al., 2020) products, at 500 and 250 m respectively. As these products are only available from November 2000 and January 2001 on, the layers are not shown in the map for earlier dates.

Both the pre-fire and the post-fire images are masked where clouds and cloud shadows could be detected. When Landsat images are used, the *pixel_qa* band is used for this purpose. With Sentinel-2 images, the *QA60* band from the L1C product is used, except when the L2A product for the same date is available; in that case, the Scene Classification or SCL is used to mask these areas.

Two new buttons are printed in the *Console* as well:

- ***Zoom to region***: a general zoom to the region.
- ***No BA observed. Export results anyway to Google Drive***: this button exports the results between the two dates for this tile, when no BA is observed and the Random Forest classifier cannot be trained; the results are exported to the user's Google Drive account. Three files are exported: the pre-fire and post-fire images (the colour compositions shown in the map) and a shapefile, which will only contain polygons for unobserved areas.
- ***No BA observed. Export results anyway as GEE asset***: the same vector layer as in the previous button is exported, but as an asset to the user's asset collection.

If BA are observed, please do not press any of these two buttons, define training polygons instead, and continue reading sections 5.3 and 5.4 below.

5.3 BA detection

Once some burned and unburned training polygons are defined using the available ***burned*** and ***unburned*** layers as *Geometry Imports* (Figure 13), burned areas can be detected for the first time. It is recommended to define polygons instead of rectangles, adjusting their shapes to burned areas, and at least one training polygon is required for each class; otherwise a new polygon will be asked to be defined.

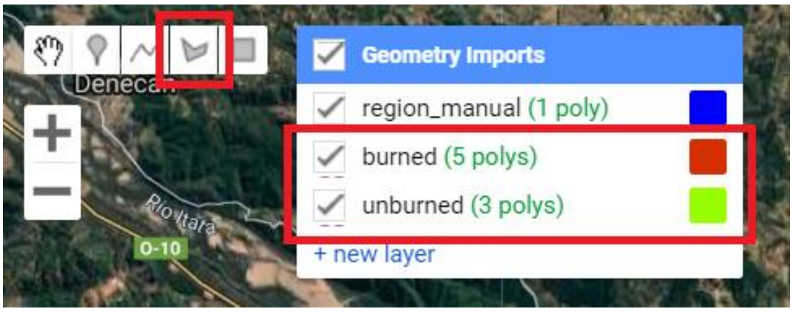


Figure 13. The layers to be used to define training polygons.

After pressing the ***Run*** button, burned areas are detected based on the training polygons. These areas are classified by a Random Forest classifier with 100 decision trees and nodes with a minimum of 10 leaves; the classifier is trained by all pixels contained in the training polygons. A total of 18 variables are introduced, the original 6 bands (blue, green, red, NIR and both SWIRs) and 3 spectral indices (NDVI, NBR and NBR2), each one of them twice: the post-fire value, and the difference between the pre- and post-fire values.

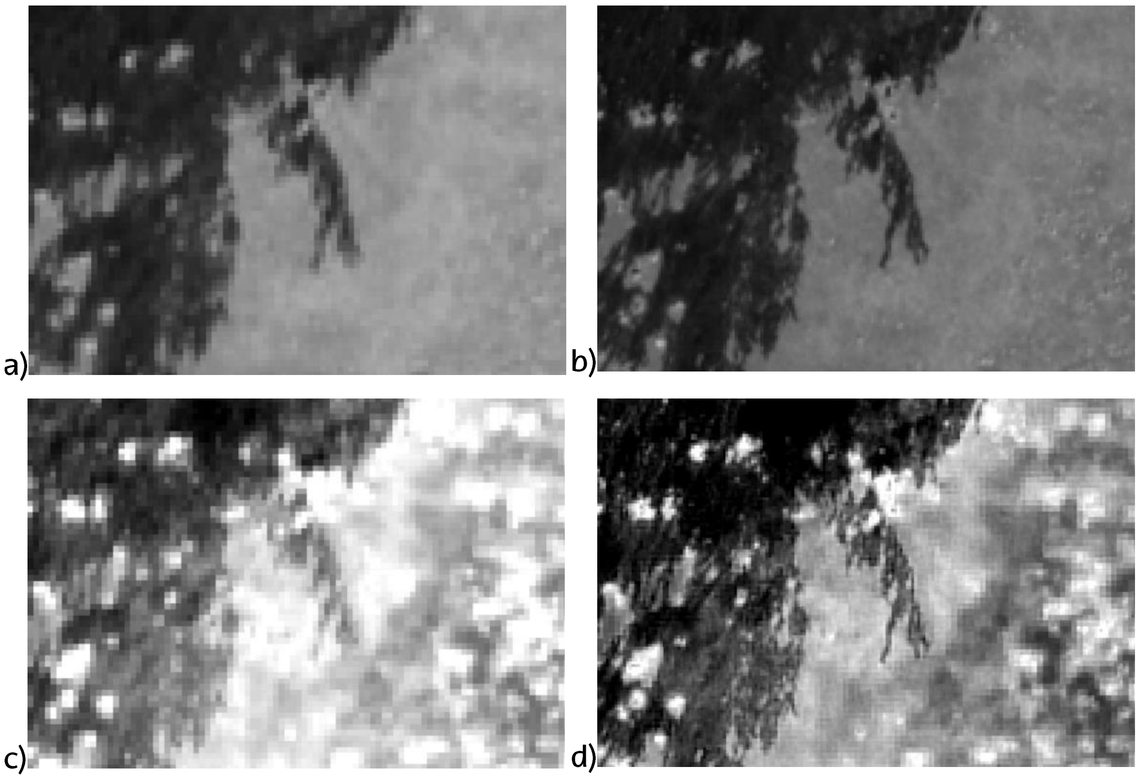


Figure 14. Sample area in S2 tile 36NVP, in South Sudan. a) NIR band at 20 m (B8A); b) NIR band at 10 m (B8); c) NBR index at 20 m (from B8A and B12); and d) NBR index at 10 m (from B8 and B12).

With S2 data, the user can select to process BA either at 20 or 10 m; the only difference is the NIR band being B8A or B8, respectively. In the first case, the original blue, green and red bands have a spatial resolution of 10 m, but 20 m for the NIR (B8A) and both SWIR bands, so BA are processed at 20m (Figure 14a and c). In the second case, the visible bands, the NIR (B8) and the NDVI index are at 10 m, and both SWIR bands and the NBR2 spectral index at 20 m. The resolution of the spectral index NBR is located halfway, as GEE can combine bands at different resolutions: NIR at 10 m and Long SWIR at 20 m, giving the impression the NBR was computed with two bands at 10 m (Figure 14b and d). Therefore, when the option at 10 m is selected, the B8 band is chosen as NIR, and BA are processed at 10 m, even though not every band/index has this resolution.

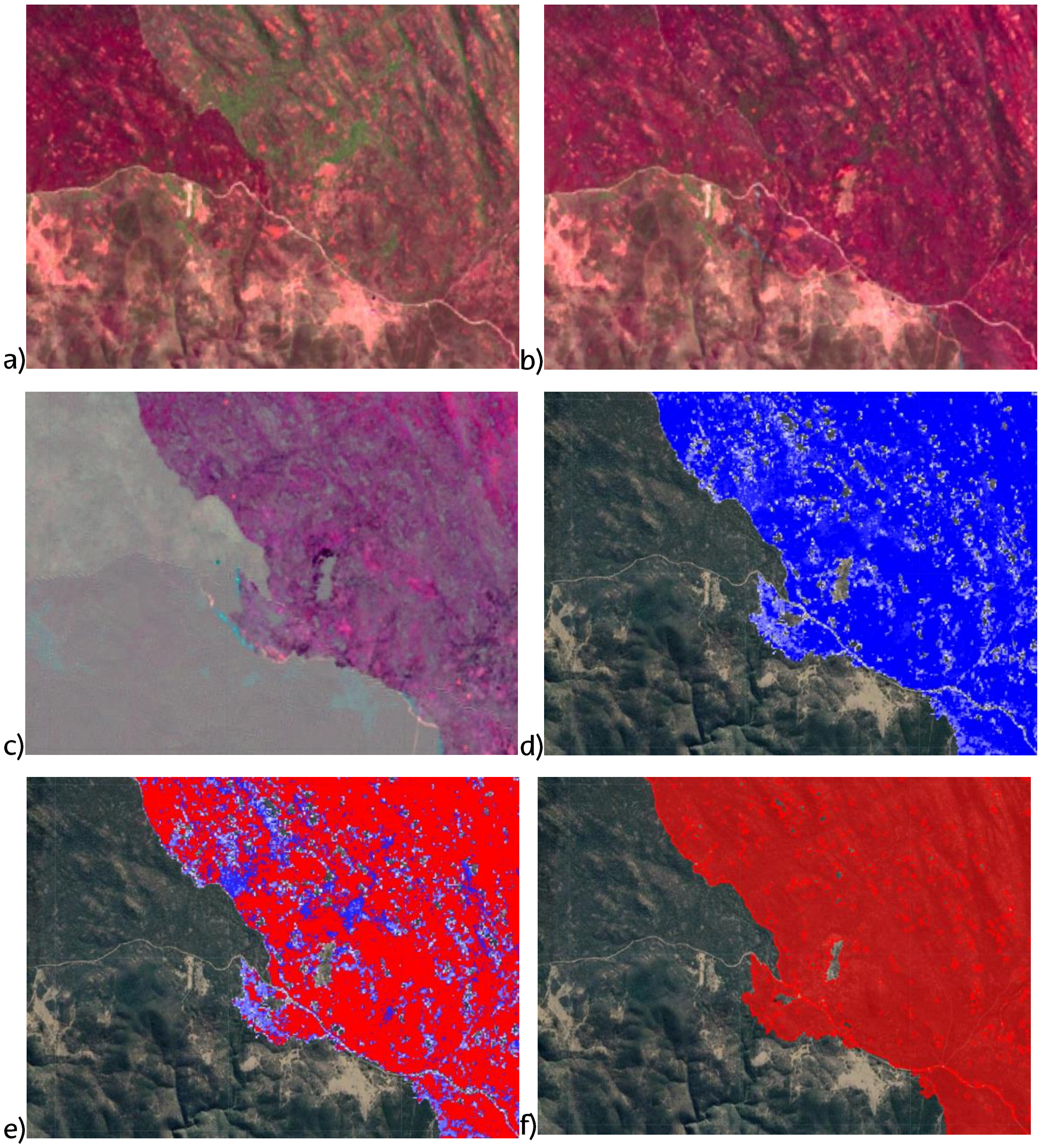


Figure 15. Sample area in S2 tile 56JLN, in Australia. a) pre-fire image from 6th November 2019, b) post-fire image from 16th November 2019, c) difference image between both dates, d) probability of burn image with the Google Satellite image in the background, e) the seeds in red shown on the previous image, and f) final vectorized result.

The result returned by the classifier is a probability image with values from 0% (unburned) to 100% (burned). Burned seeds are extracted from this image, representing pixels with the strongest burned signal; the threshold used here is the minimum among the mean burned probabilities in every burned training polygon. Three new layers are now added to the *Map* (Figure 15):

- ***Prob.***: the probability image. To make easier its visualisation, a threshold of 50% is applied to this image before being displayed, so that pixels below this threshold are transparent and the image at the bottom can be seen. Burned pixels are displayed with a colour palette from white (probability values close to 50%) to blue (100% probability of burned).
- ***Seeds***: the seeds extracted from the previous probability image in a red colour.
- ***BA***: final result, where BA are are expanded from burned seeds outwards up to the 50% threshold, so burned patches not containing any seed have already been removed. This final result is a vectorized layer, so it may take some time for GEE to display it.

The user can now modify or define new training polygons until the desired results are obtained, pressing the ***Run*** button at each iteration to update the BA.

5.4 Exporting results

When the desired result is finally reached, it can be exported to the user's Google Drive account by the ***Export BA to Google Drive*** button in the *Console*. A total of 4 different files can be downloaded: the 1) pre-fire and 2) post-fire images with the colour compositions shown in the map, 3) the probability of burn image (from 0 to 100%), and 4) the shapefile with polygons. These polygons have only one attribute, with a *BURNED* or *UNOBSERVED* label; observed and unburned areas are not included in the shapefile. Except for the probability of burn image, the files are identical to those mentioned at the end of section 5.2.

The files have the following names:

1) **BAMT_RP_*ID*_*SENSOR*_TILE-*TILEID*_*DATE1*_RGB.tif**
2) **BAMT_RP_*ID*_*SENSOR*_TILE-*TILEID*_*DATE2*_RGB.tif**
3) **BAMT_RP_*ID*_*SENSOR*_TILE-*TILEID*_*DATE1*_*DATE2*_TIF.tif**
4) **BAMT_RP_*ID*_*SENSOR*_TILE-*TILEID*_*DATE1*_*DATE2*_SHP.shp**

where:

- **ID**: the identifier defined by the user.
- **SENSOR**: *S2MSI* or *Lndst*, depending on the used dataset, corresponding to Sentinel-2 MSI and Landsat respectively.
- **TILEID**: the ID of the tile depends on the selected value for the *region_type* variable:
    - Landsat: ***DIM*km-*PPPRRR***, where ***DIM*** stands for the width and height of the square in the centre of the Landsat scene, and ***PPP*** and ***RRR*** stand for the path and row respectively, each with three digits. E.g. '50km-090080' for a square 50 km wide and high in the centre of the Landsat scene with path 90 and row 80.
    - Sentinel: ***DIM*km-*TTTTT***, where ***DIM*** stands for the width and height of the square in the centre of the Sentinel tile, with ***TTTTT*** being the tile's code. E.g. '50km-56JLN' for a square 50 km wide and high in the centre of the S2 56JLN tile.
    - Manual: **Mnl-*LAT*-*LON***, where the **LAT** and **LON** represent the latitude and longitude of the centre of the region defined by the user, with one decimal and the Northern/Southern and Western/Eastern hemisphere suffixes. E.g. 'Mnl-29.4S-151.5E'.
- **DATE1**: the pre-fire date, in *YYYYMMDD* format.
- **DATE2**: the post-fire date, in *YYYYMMDD* format.

The whole final BA result can be exported as a GEE asset, instead of an ESRI Shapefile to the user's Google Drive account, with the button ***Export BA as GEE asset***; this way, the result is stored in a folder called *BAMT_RP*. This option is recommended when the results are intended to be used for for a validation process with the Assessment tool; otherwise, all ESRI Shapefiles exported to Google Drive will have to be uploaded to GEE instead.

Finally, once the BA detected and all the necessary data exported, it is important not to forget to delete some polygons before beginning to detect BA elsewhere. The *region_manual* layer cannot have more than a rectangle at the same time, so when the option *manual* is selected as region type the previous rectangle should be removed before changing the region. As for burned and unburned training polygons, the tool filters these polygons with the region, so it is not necessary to delete them when changing region. They must be deleted, however, if the dates are changed but not the region, as pixels in training polygons will hardly correspond to the same class.

However, the training polygons can be saved in case the process must be repeated. It is as simple as saving a copy of the script. While the user continues working in the original script, the region, the dates and the training polygons will be kept safe in the copy.

6 Assessment tool

https://code.earthengine.google.com/33664838272530f4b1a3a199fcd954bd

This last tool compares a BA product with reference data and carries out its validation, returning the necessary accuracy metrics. The BA product can be a global BA product already available in GEE, such as MCD64A1 or FireCCI51, or the resulting BA maps created with the BA Cartography tool. The user does not have to handle any code, since the entire tool works through a simple interface (Figure 16). However, both the BA maps and the reference data must follow the format described here below for the tool to correctly interpret and manage them.

Figure 16. Example image of the *Assessment* tool and its interface; in the map, a BA product is validated in a validation site.

6.1 Initial parameters

Every file that is to be used by this tool needs to be uploaded into GEE as an asset with an ID of the project in its name; this ID must be defined in the ***Project identifier*** parameter, and will be used by this tool to distinguish layers contained in the same folders but corresponding to other projects.

6.1.1 RP

Reference data may consist of perimeters generated between just one pair of images in a validation site (short unit), or in a series of consecutive pairs (long unit). If the user employed the RP tool to create these perimeters between several pairs of images, the files are in the correct format and only need to be uploaded to the asset collection in GEE; if the perimeters were directly exported as assets, uploading them into GEE is not even necessary, since they are already in a folder called *BAMT_RP* in the asset collection (Figure 17).

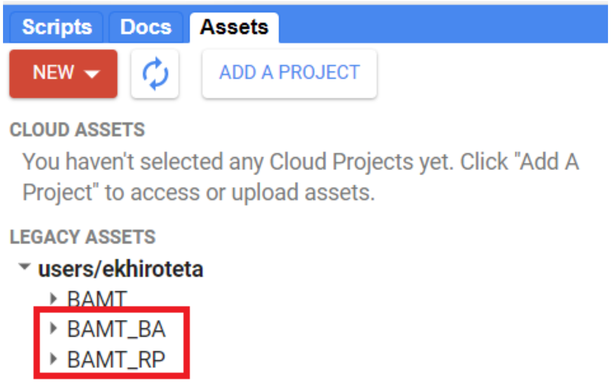


Figure 17. Location of the *BAMT_BA* and *BAMT_RP* folders, where exports by the BA Cartography and RP tools are stored, in the user's asset collection.

However, if the user wishes to use reference data created by other means than the RP tool, first they have to be adapted to the same specific format. For vector layers, the assets' names must follow the format already described in section 5.4 (BAMT_RP_*ID*_*SENSOR*_TILE-*TILEID*_*DATE1*_*DATE2*_SHP), and they must have a column called *LABEL* with an attribute of *BURNED* or *UNOBSERVED* for each polygon; observed but unburned polygons should not be represented in these layers. If raster layers containing reference data are uploaded instead, the final element of the asset should be *_TIF* instead of *_SHP*, and the pixel values must be either 0 (unburned), 1 (burned) or 2 (unobserved).

Once reference data from all short units are uploaded as assets in GEE, they must be merged into one final layer, corresponding to the long unit in that validation site. This new layer displays which pixels burned during the whole validation period, as it is shown in Figure 18, as well as labelling pixels that were not observed at least in one image of the series with an "unobserved" label. These long unit layers are created with the button ***Merge short units***, available in the tool's interface. This button checks all the assets contained in the indicated folder, classifies them among different validation sites, and merges all the short units at each site. Each new layer corresponding to a long unit, one per validation site, is displayed in the *Map* and appears in the *Tasks* tab; the user must ***Run*** every task in order to process them. After each task is completed, the new long unit layer will be stored as an asset, with *_LU_TIF* added to its asset name.

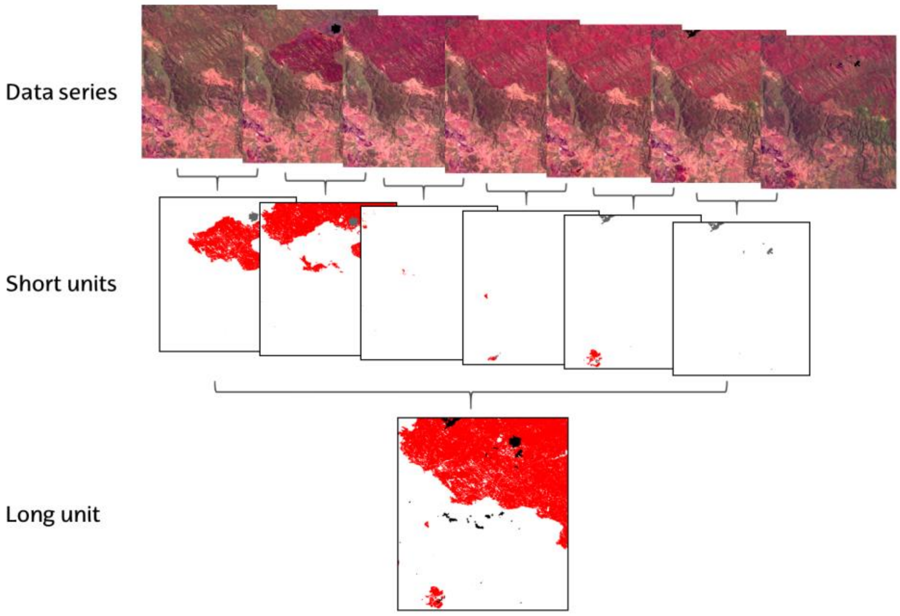


Figure 18. Generation of a long unit reference data layer by merging several short units.

The same long unit image can also be uploaded by the user to GEE. In that case, it must be made sure that the uploaded asset is called with the same name format (BAMT_RP_*ID*_*SENSOR*_TILE-*TILEID*_*DATE1*_*DATE*2_LU_TIF), and that pixels have exactly the same values (0 – unburned, 1 – burned and 2 – unobserved).

6.1.2 BA product

Two types of BA products are accepted by this tool: vector layers and an image collection. Vector layers must follow the same format as those obtained with the BA Cartography tool. If they were directly exported as assets by the first BAMT tool, they are already ready to be assessed in a folder called *BAMT_BA* (Figure 17); otherwise, the shapefiles will have to be uploaded to GEE. Each asset's name must end in *_SHP*, as well as having the ID of the project in the sixth-from-last position and delimited by underscores (*_ID_*). All these layers must also have a column called *BurnDate*, with a value of YYYYMMDD for every polygon, where YYYY, MM and DD are the year, the month and day of the month where the burning was registered; if unobserved areas such as clouds or water bodies are represented, their *BurnDate* values should be lower than 19,000,000 so that they cannot be mistaken for a date.

The other type of accepted BA product is an image collection, consisting in a series of images corresponding to monthly BA products. Each monthly image may have multiple bands, but only one is required: a band called *BurnDate*. This band must indicate for every burned pixel the day of the year from 1 to 365 (or to 366 in the case of gap years) where they were detected, or a value outside that range (lower than 1, or higher than 366) for the rest of the pixels (unburned, cloudy or water bodies). Each monthly image must also contain an attribute called *system:index* indicating the first day of the corresponding month in a YYYY_MM_DD format, as well as two attributes called *system:time_start* and *system:time_end* representing the first day of the corresponding and next months, respectively; the values of these last two attributes are the number of milliseconds since the beginning of the year 1970, which GEE needs in order to filter the image collection by date. This is the exact format of BA products already available in GEE (MCD64A1 and FireCCI51), so the tool can also be used in order to assess these global products.

6.2 Assessment

Once both the reference perimeters and the BA product are adequately defined, the assessment of the product can be carried out. The user must specify if it should be based on short or long units (Figure 18). If short units are chosen, the BA product is assessed individually between every pair of images, and this tends to show higher errors due to temporal differences on BA detections, so long units are recommended instead.

After pressing the button ***ASSESS***, the validation of the BA product is carried out at a scale of 10 m. The results of each validation site can be seen in the images displayed in the *Map*; each one of these images shows which areas were committed, omitted, or detected correctly (Figure 16). A list of features is also shown in the *Console* of GEE, with one feature per each validation site, and another last feature with aggregated validation results of the whole BA product. Each feature contains the following attributes:

- ***TILE***: an ID of the validation site, corresponding to the MGRS tile name or the path and row for Sentinel-2 and Landsat data, respectively. For the aggregated results of the whole BA product, a value of *TOTAL* is used in this attribute.
- ***DATE1***: beginning of the validation period.

- ***DATE2***: end of the validation period.
- ***surfBA***: total surface correctly detected as burned in km^2.
- ***surfUB***: total surface correctly detected as unburned in km^2.
- ***surfaCE***: total committed burned surface in km^2.
- ***surfOE***: total omitted burned surface in km^2.

As well as these accuracy metrics based on the confusion matrix:

- ***OA***: overall accuracy.
- ***CE***: commission error.
- ***OE***: omission error.
- ***Kappa***: the Kappa coefficient (Congalton and Green, 2008).
- ***relB***: relative bias, which indicates if BA are underestimated or overestimated by the BA product.
- ***DC***: the Dice coefficient (Dice, 1945) which, given two BA classifiers, indicates the probability for one classifier to detect a pixel as burned if the other classifier detected it as burned, too.

These results can be exported to an ESRI Shapefile, which will be stored in the user's Google Drive account, in a folder called *BAMT_GEE*. The file will be named this way:

1) **BAMT_AS_*ID*_*UN*_SHP.shp**

where:

- **ID**: the project's identifier as defined by the user.
- **UN**: *SU* or *LU*, depending on whether the assessment is based on short or long units, respectively.

At the same time, the resulting validation images representing committed, omitted and correctly detected areas can be exported (Figure 16) as well, if the option *Export validation images* is checked. The name of each exported image follows this naming format, which is practically the same as that exported by the RP tool:

1) **BAMT_AS_*ID*_*SENSOR*_TILE-*TILEID*_*DATE1*_*DATE2*_*UN*_TIF.tif**

where:

- **ID**: the identifier defined by the user.

- **SENSOR**: *S2MSI* or *Lndst*, depending on the used dataset, corresponding to Sentinel-2 MSI and Landsat respectively.
- **TILEID**: the ID of the tile is the same as in the reference image files, and is defined in section 5.4.
- **DATE1**: the pre-fire date of the short or long unit, in *YYYYMMDD* format.
- **DATE2**: the post-fire date of the short or long unit, in *YYYYMMDD* format.
- **UN**: *SU* or *LU*, depending on whether the assessment is based on short or long units, respectively.

The possible pixel values of these raster files are:

- **0**: unburned.
- **1**: omitted.
- **2**: committed.
- **3**: burned.
- **4**: unobserved areas according to the reference data (NoData, clouds, water bodies, etc.).

7 Image Viewer tool

https://code.earthengine.google.com/478e4218e6fc00d51f9b36e49073950a

This is an extra tool, not related to BA mapping, which serves for the user to view Landsat or Sentinel-2 images throughout the world. The tool is very easy to handle, since it works via an interface just as the VA Dates and Assessment tools. The interface appears after pressing the button ***Run***, with these parameters:

1) ***Dataset***:the data must be chosen from a list that contains all the possible options:
   - ***Sentinel-2 MSI L1C***: L1C products of the MSI sensor aboard Sentinel-2A and 2B satellites.
   - ***Sentinel-2 MSI L2A***: L2A products of the MSI sensor aboard Sentinel-2A and 2B satellites.
   - ***Landsat-9 OLI-2***: OLI-2 sensor aboard the Landsat-9 satellite.
   - ***Landsat-8 OLI***: OLI sensor aboard the Landsat-8 satellite.
   - ***Landsat-7 ETM+***: ETM+ sensor aboard the Landsat-7 satellite.
   - ***Landsat-5 TM***: TM sensor aboard Landsat-5.
   - ***Landsat-4 TM***: TM sensor aboard Landsat-4.
   - ***Landsat 4-9***: sensors TM, ETM+, OLI and OLI-2 aboard satellites Landsat-4 to 9.
2) ***Region***: if Sentinel-2 data are selected, the user must choose between options ***Sentinel-2 Tile*** and ***Map view***; if some Landsat dataset is selected, between ***Landsat Scene*** and ***Map view***.
   - ***Sentinel-2 Tile***: the images are filtered by the S2 ***Tile*** defined by the user.
   - ***Landsat Scene***: the images are filtered by the Landsat ***Path*** and ***Row*** defined by the user.

     There are two layers called ***Sentinel-2 tiles*** and ***Landsat scenes*** available in the map for the user to consult, in case the ID of the S2 tile or the path/row of the Landsat scene are unknown. This information will appear in the *Inspector* tab after clicking on them in the *Map*.
   - ***Map view***: the displayed images will be those covering the area of the current map view.
3) ***Dates***: the beginning and the end of the period to display.

4) ***Cloud coverage***: images with a cloud coverage higher than this value will not be displayed. If ***Sentinel-2 Tile*** or ***Landsat Scene*** was selected as region, images are filtered with the cloudiness value in the metadata. Otherwise (***Map view***), the cloud coverage is computed for every image in the *Map view* area; for large areas or long periods, this may slow down or even crash the process with a message of '*Too many concurrent aggregations*' printed in the *Console*.
5) ***Visualization parameters***: here the colour composition is formed by selecting the bands for each RGB colours, from a list with 6 bands (Blue, Green, Red, NIR, Short SWIR or SWIR1, and Long SWIR or SWIR2), and 4 spectral indices (NDVI, NBR, NBR2 and MIRBI). All bands can be stretched between minimum and maximum values defined on the right.
   Four predefined colour compositions are already available for the user: ***Natural colour*** (Red-Green-Blue), ***Modified natural colour***, ***NIR-Red-Green*** and ***SWIR2-NIR-Red***. The ***Modified natural colour*** composition tries to remove the effect of the atmospheric dispersion to the ***Natural colour***, so it is recommended for the ***Sentinel-2 MSI L1C*** dataset, which contains Top-Of-Atmosphere reflectances.
   Finally, a temporal composite can be created by checking the checkbox right to ***Composite***, thus maximazing (***Max***) or minimazing (***Min***) the selected band or spectral index. If ***Mask*** is selected instead of ***No mask***, clouds and cloud shadows are masked before maximazing or minimazing.

After defining all these parameters, the button ***Load and display*** can be pressed, and all the available images in the selected region for the defined period will be displayed in the *Map*, all of them with a cloud coverage lower than the defined threshold; the date of each image can be seen in the *Layers* list in the *YYYYMMDD* format (Figure 19).

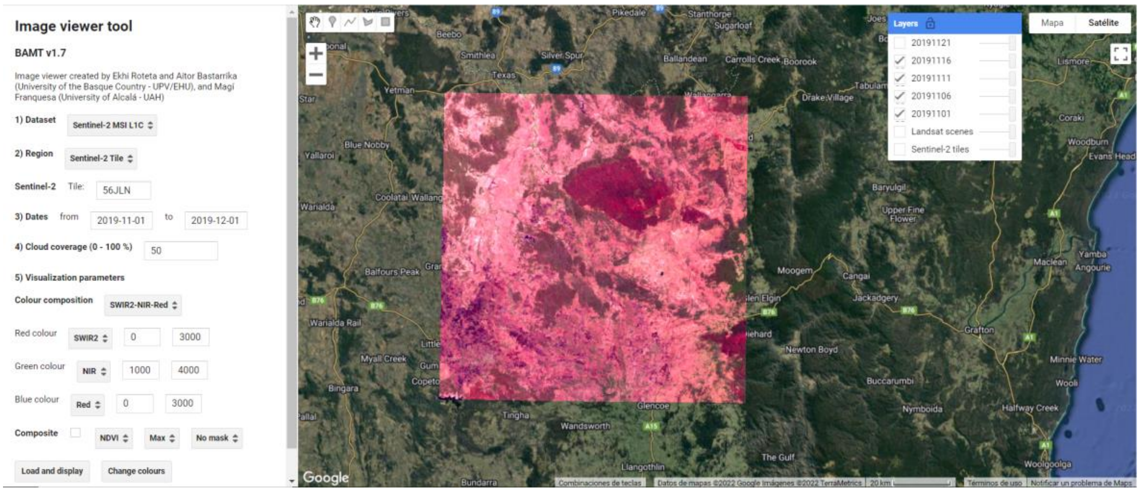


Figure 19. Example image of the *Image viewer* tool; input parameters can be seen in the interface on the left, and the resulting images in the map.

A message telling the number of found images is printed below the buttons; if no image was found in the area for the defined period, the message is '*No image found. Please change parameters*' instead.

Once the images have been displayed, a new button is available: ***Change colours***. This button is recommended whenever the user wants to change ***Visualization parameters***, but keeping the dataset, region, dates and cloud coverage; since the images are already loaded, this button just changes the visualization parameters, without loading images again. This way the process is faster, especially if ***Map view*** was selected as region.

8 Updates from previous versions

This version 1.7 of the BAMT implies several improvements when compared to its predessor BAMT v1.6. Some of these improvements updated on every tool are:

- Inclusion of Landsat-9 data, which were not available yet when the previous version 1.6 was made public (the Landsat-9 satellite was launched in September 2021). The sensor aboard is the same as that on the Landsat-8 satellite, despite being called slightly different names (OLI vs. OLI-2).
- Replacement of the old LSR dataset in the GEE catalog (https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C01_T1_SR, last accessed in April 2022), which is now deprecated, by the new dataset (https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C02_T1_L2, last accessed in April 2022) for all Landsat satellites.
- Replacement of the previous S2 dataset by harmonized Sentinel-2 data. According to GEE, the DN values of S2 images from January 25, 2022 onwards are shifted by 1,000 and they cannot be directly compared with previous images, so a new *harmonized* collection has been created to be used instead (https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_HARMONIZED, last accessed in April 2022).
- Replacement of the method *.filterMetadata()*, which is still working but deprecated, by the corresponding filter as *.filter(ee.Filter...)*.

Some other improvements are related to specific tools:

- BA Cartography and RP tools allow exporting their results as assets to the user's asset collection, along with the regular export to the Google Drive account. This way using the new Assessment tool is more straightforward, since the BA product and the reference perimeters do not need to be uploaded one by one.
- The same final result that will be later exported can be seen in the *Map* with the RP tool; only the probability image by the RF classifier and the burned seeds were displayed in the previous version, and the effect of extending BA around seeds could not be seen until the results were exported. This

cannot be done for the BA Cartography tool yet, since the study area is too large for GEE to vectorize the whole result.

- The whole process for selecting validation sites with the stratified random sampling methodology (VA tool) has been optimized, so now it is able to process larger study areas.
- Enabling the sampling units to be exported in the same VA tool. This layer was displayed in the GEE interface by the previous version, but could not be exported to the user's Google Drive account.

Finally, the Assessment tool is released for the first time in this version v1.7, allowing the validation of a product to be carried out within the GEE environment. Previously, it had to be designed and conducted by the user in a local system.

9 References

Bastarrika, A., Alvarado, M., Artano, K., Martinez, M., Mesanza, A., Torre, L., Ramo, R., Chuvieco, E., 2014. BAMS: A Tool for Supervised Burned Area Mapping Using Landsat Data. Remote Sens. 6, 12360–12380. https://doi.org/10.3390/rs61212360

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

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

Chuvieco, E., Lizundia-Loiola, J., Lucrecia Pettinari, M., Ramo, R., Padilla, M., Tansey, K., Mouillot, F., Laurent, P., Storm, T., Heil, A., Plummer, S., 2018. 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. https://doi.org/10.5194/essd-10-2015-2018

Cohen, W.B., Yang, Z., Kennedy, R., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync - Tools for calibration and validation. Remote Sens. Environ. 114, 2911–2924. https://doi.org/10.1016/j.rse.2010.07.010

Congalton, R.G., Green, K., 2008. Assessing the Accuracy of Remotely Sensed Data, Assessing the Accuracy of Remotely Sensed Data. https://doi.org/10.1201/9781420055139

Dice, L.R., 1945. Measures of the Amount of Ecologic Association Between Species. Ecology 26, 297–302. https://doi.org/10.2307/1932409

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. https://doi.org/10.5194/essd-12-3229-2020

García, M.J.L., Caselles, V., 1991. Mapping burns and natural reforestation using thematic Mapper data. Geocarto Int. 6, 31–37. https://doi.org/10.1080/10106049109354290

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

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 Sens. Environ. 113, 408–420. https://doi.org/10.1016/j.rse.2008.10.006

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/https://doi.org/10.1016/j.rse.2017.06.031

Kennedy, R.E., Yang, Z., Cohen, W.B., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms. Remote Sens. Environ. 114, 2897–2910. https://doi.org/10.1016/j.rse.2010.07.008

Key, C.H., Benson, N., 1999. The Normalized Burn Ratio (NBR): a Landsat TM radiometric measure of burn severity. US Geol. Surv. North. Rocky Mt. Sci. Center.

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

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., Kassem, K.R., 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. Bioscience 51, 933. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Padilla, M., Olofsson, P., Stehman, S. V., Tansey, K., Chuvieco, E., 2017. Stratification and sample allocation for reference burned area data. Remote Sens. Environ.

203, 240–255. https://doi.org/10.1016/j.rse.2017.06.041

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

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., 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/10.1016/j.rse.2015.01.005

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

Rouse, J.W., Haas, R.H., Schell, J.A., 1974. Monitoring the vernal advancement and retrogradation (greenwave effect) of natural vegetation. NASA Goddard Sp. Flight Cent. https://doi.org/19740008955