Final Report

Reporting Period: July 1\textsuperscript{st} 2015 to June 30\textsuperscript{th} 2018

Prototyping a Landsat-8 Sentinel-2 Global Burned Area Product

PI: David Roy

Geospatial Sciences Center of Excellence, South Dakota State University, USA

NASA NNH14ZDA001N

\textit{Land Cover/Land Use Change (LCLUC14-2): Multi-Source Land Imaging Science.}

1.0 Proposal Summary

This proposal prototyped a global burned area product by combination of NASA-USGS Landsat-8 and ESA Sentinel-2 data and was directly responsive to the primary focus of the call - developing algorithms and prototyping products for combined use of data from Landsat-8 and Sentinel-2 toward global land monitoring, and advances the virtual constellation paradigm for mid-resolution land imaging. Product quality assessment and validation was conducted by comparison with the MODIS Collection 6 burned areas product and with visually interpreted commercial high resolution PLANET data. The advantages of the proposed sensor combination are (i) Landsat-8 has improved quantization, signal/noise characteristics, and acquisition coverage over heritage Landsat missions, (ii) Sentinel-2 has Landsat-8 like bands at 10 m and 20 m with higher acquisition coverage than Landsat-8, (iii) Landsat-8 and Sentinel-2A/B together provide the needed temporal resolution for time series burn change detection.
2.0 Research progress executive summary

This has been a productive grant and resulted in 17 published papers (see Section 7). Despite early unexpected difficulties obtaining and using the Sentinel-2 data, an end-to-end Landsat-8 and Sentinel-2 processing chain, and an automated sensor-agnostic burned area mapping algorithm, were developed. The African 30 m burned areas results are very encouraging when compared to the most recent Collection 6 MODIS burned area product (see Section 5).

In **Year 1** the research emphasis was on

- obtaining and understanding the Sentinel-2 Multi Spectral Instrument (MSI) data format and developing Sentinel-2 processing under the WELD processing framework
- developing an automated Landsat-8 to Sentinel-2 registration methodology

In **Year 2** the research emphasis was on:

- refining Landsat-8 and Sentinel-2 processing under the WELD processing framework, including dealing with Landsat-8 and Sentinel-2 format changes, and BRDF, registration, and atmospheric corrections
- refining the automated sensor-agnostic burned area algorithm
- developing burned area product comparison and validation approaches

In **Year 3** the research emphasis was on:

- running the sensor-agnostic burned area algorithm
- comparison of the results with the MODIS Collection 6 burned area product

3.0 Development of Landsat and Sentinel-2 processing

3.1 Obtaining and understanding the Sentinel-2 data and format and reprojection with Landsat-8 data to common projection and tiled grid

Due to ground system limitations, the Sentinel-2 data coverage in the early post-launch period was limited predominantly to Europe and Africa and only a limited number of pre-operational Sentinel-2 product samples were publicly available. In late November 2015 the ESA publicly released the operational data products. A script to pull the Sentinel-2 L1C data acquired over Africa from the Sentinels Scientific Data Hub was developed to pull large product volumes. Code was developed to read the Sentinel-2 data and reproject it into the Global WELD format compatible with Landsat-8 WELD processed data.

The Sentinel-2 data are made available as L1B and L1C top-of-atmosphere (TOA) products. The L1B products are defined in the sensor swath geometry and are not currently available. The L1C products are defined in reflectance units and include cloud and land/water masks that are not provided in the L1B product. The geolocated L1C products are defined by splitting each MSI swath into fixed 109 × 109 km tiles in the Universal Transverse Mercator (UTM) map projection. The L1C data are provided in Standard Archive Format for Europe (SAFE) files.
Code to read the L1C data format and extract the relevant 10m and 20m bands was written. The L1C tiled data structure is quite complicated for users to handle without reliance on dedicated software tool kits. In particular, adjacent tiles in the same MSI swath may overlap spatially, and may be defined in different UTM zones. The UTM projection is defined in zones, each covers 6° of longitude centered over a meridian of longitude, and so each zone forms the basis of a separate map projection. A letter was published in Remote Sensing Letters (Roy et al. 2016) to (a) illustrate the spatial properties of Sentinel-2 L1C data and provide insights into the occurrence of overlapping tiles and overlapping tiles defined in different UTM zones, (b) quantify the geometric implications of resampling and reprojection approaches that consider only the data from one tile and not the data from other tiles in the overlap region that are defined in different UTM zones, and (c) suggest and illustrate a recommended approach for resampling and reprojection of Sentinel-2 L1C data.

We found that Sentinel-2 L1C resampling and reprojection approaches that consider only the data from one L1C tile, and not from other tiles in the overlap region defined in different UTM zones, will result in a significant degradation of the geometric fidelity of the resampled image. This is illustrated in Figure 1 - the linear high contrast features are more coherent and the geometric fidelity is improved when the recommended method is used. The geometric differences among these three sets of images are greatest comparing the results derived from the separate L1C zone tiles, i.e., comparing the (a) and (b) results. The recommended approach considers both sets of L1C data and therefore a greater density of Sentinel-2 observations is available for resampling which results in the improved geometric fidelity evident in Figure 1(c). A straightforward scheme is to process each Sentinel-2 L1C tile independently in sequence was developed. For each resampled pixel location the distance to the closest Sentinel-2 pixel is stored. Each time a resampled pixel location is projected into a Sentinel-2 tile its distance to the closest Sentinel-2 pixel is derived and the Sentinel-2 pixel data are selected only if the distance is smaller than the distance found for the previous L1C tile. In this way only one Sentinel-2 L1C tile at a time needs to be stored in memory, each UTM zone is treated separately, and the resampling results are the same regardless of the number or processing order of the overlapping Sentinel-2 L1C tiles.

![Figure 1. Example nearest neighbor resampled Sentinel-2 10m L1C tile data (250 × 250 pixels) from the same Sudan SAFE file (i.e., MSI swath) considering (a) only one L1C tile defined in UTM zone 35N, (b) only one L1C tile defined in UTM zone 36N, (c) both overlapping tiles using the recommended resampling method. True color 665nm (red), 560nm (green), 490nm (blue) top of atmosphere Sentinel-2 reflectance.](image)
Code to reproject the Sentinel-2 L1C data format to the global WELD projection and tiling scheme was developed. Over large areas, users of derived satellite information require spatially explicit map products defined in a single projection rather than different UTM projections. Therefore, in this study, the Sentinel-2 L1C and Landsat-8 L1T data were reprojected to a common coordinate system. The equal area sinusoidal projection and tiling scheme used to store the global Web Enabled Landsat (WELD) products were used. The global WELD products define monthly and annual 30m Landsat nadir BRDF-adjusted reflectance (NBAR) surface reflectance and are available at [http://globalweld.cr.usgs.gov/collections/](http://globalweld.cr.usgs.gov/collections/). The global WELD tiles are nested within the standard 10° × 10° MODIS land product tiles and are defined in the sinusoidal equal projection used to store the MODIS land products. Each global WELD tile covers about 159 × 159 km. There are 7 × 7 global WELD tiles within each MODIS land tile, and the filename includes the MODIS horizontal (0 to 35) and vertical (0 to 17) tile coordinates, and the nested WELD tile horizontal and vertical tile coordinates (0 to 6).

Code to resample the Sentinel-2 L1C data were to 30m output WELD tile pixels was developed. Only the 20m and 10m Sentinel-2 bands are used for burned area mapping and they were projected using nearest neighbor and box-car resampling respectively (Fig. 2). In addition, a code to reproject the near-infrared 10m Sentinel-2 L1C data to 10m output WELD tile pixels was developed to support the Sentinel-2 to Landsat-9 misregistration code.

Figure 2. Landsat 8 (L8) and Sentinel 2 (S2) have different spectral & spatial resolutions. Illustration of how the different bands are resampled to 159 × 159 km global WELD tiles.
The reproject and tiling scheme was refined to accommodate new data formats. The Sentinel-2 data format was changed by ESA to store L1C tiles in separate files, rather than bundled into ESA SAFE files, and provided with a shorter file name. The USGS Landsat Collection 1 Landsat products became available May 2017 for all the Landsat 4-8 archive. They have more rational filenames and provide top-of-atmosphere (TOA) radiance/reflectance with improved per-pixel cloud masks, quality data, calibration information, and new metadata enabling per-pixel view and solar geometry calculation.

3.2 Registration correction

The Sentinel-2 geolocation uses a Global Reference Image (GRI) derived from orthorectified Sentinel-2 cloud-free images that is scheduled for completion by the last quarter of 2017. The Landsat framework will be readjusted for consistency with the Sentinel-2 GRI, with completion expected in 2018 (Storey et al. 2016), leading to a complete reprocessing of the Landsat archive as Collection 2. When the data in overlapping Sentinel-2A swath edges were compared, the overlapping data were found to be misregistered by >30 m due mainly to a satellite yaw orientation knowledge error that was not rectified in the Sentinel-2A processing until Summer 2016 (ESA 2016). Given the uncertainty about the Sentinel-2A reprocessing schedule, we developed an algorithm to fix the registration of the Sentinel-2 and Landsat-8 data.

The registration algorithm uses a hierarchical image pyramid approach with area- and feature-based matching and mismatch detection and handles translational, rotational and scale differences to provide sub-pixel registration accuracy (Yan et al. 2016). It was based on an algorithm developed for registration of High-Resolution Imaging Science Experiment (HiRISE) single-band stereo images to derive digital elevation models of Mars. The approach is computationally efficient because it implements feature point detection at reduced spatial resolution and then area-based least squares matching around the feature points with mismatch detection across four image pyramid levels to identify a sparse set of tie-points. It was assessed by examination of extracted tie-point spatial distributions and tie-point mapping transformations (translation, affine and polynomial), dense matching prediction-error assessment, and by visual registration assessment.

Test sites over Cape Town (Fig. 3) and Limpopo province in South Africa that contained cloud and shadows were selected. A Landsat-8 L1T image and two Sentinel-2 L1C images sensed 16 and 26 days later were registered (Cape Town) to examine the robustness of the algorithm to surface, atmosphere and cloud changes, in addition to a Landsat-8 L1T and Sentinel-2 L1C image pair sensed 4 days apart (Limpopo province). The automatically extracted tie-points revealed sensor misregistration greater than one 30 m Landsat-8 pixel dimension for the two Cape Town image pairs (Fig. 4), and greater than one 10 m Sentinel-2 pixel dimension for the Limpopo image pair. Transformation fitting assessments showed that the misregistration can be effectively characterized by an affine transformation. Hundreds of automatically located tie-points were extracted from the registered sensor pairs and had affine transformations with root mean square error fits of 0.3 10m pixels and dense matching prediction-errors of similar magnitude. This and visual assessment of the affine transformed data indicate that the methodology provides sub-pixel registration performance required for meaningful Landsat-8 OLI and Sentinel-2 MSI data comparison and combined data applications. The algorithm was published in a Sentinel-2 special edition of Remote Sensing (Yan et al. 2016).
Figure 3. Cape Town, South Africa, test data showing (a) Landsat-8-L1T sensed November 22nd 2015 (week 47), (b) Sentinel-2 L1C sensed December 8th 2015 (week 49), (c) Sentinel-2 L1C sensed December 18th 2015 (week 51). The NIR (Sentinel-2: 842 nm and Landsat-8 864nm band) TOA reflectance for each image is shown reprojected to 10 m global WELD tile hh19vv12.h3v2 (sinusoidal projection, 15885 × 15885 10 m pixels).
Figure 4. Illustration of the tie-points and misregistration vectors obtained from the Cape Town data (Figure 3). The vectors point from the Landsat-8 week 47 image tie-point locations to the Sentinel-2 week 49 (116 green vectors) and to the Sentinel-2 week 51 (797 red vectors) tie-point locations. The vector lengths are enlarged by 80 times for visual clarity. To provide geographic context, the background image shows the Landsat-8 week 47 30 m true color image. For both image pairs, the x-axis and y-axis mean shift magnitudes are greater than 5.2 and 2.1 pixels with standard deviations of about 0.4 and 0.3 pixels, respectively. These 10 m pixel shifts are not insignificant and even at the 30 m Landsat pixel resolution will limit the ability to meaningfully compare Landsat-8 and Sentinel-2 data.

The registration algorithm was adapted to work on orbits of data and is described in (Yan et al. 2018). It was found that the Sentinel-2A tile-to-tile mean misregistration was 1.6 10m pixels before June 15th 2016 when ESA updated the processing software to version 2.04; after the update, the mean misregistration was reduced to 0.4 10m pixels. The new orbit-based registration method provided mean misregistration less than 0.15 10m pixels.

A public open-source software version was developed under an unsolicited proposal to NASA and is available at https://openprairie.sdstate.edu/landsat_sentinel_registration/2/.
3.4 Atmospheric correction

Surface reflectance, i.e., top of atmosphere (TOA) reflectance corrected for atmospheric effects, is needed because the impact of atmospheric gases and aerosols on reflective wavelength radiation is variable in space and time. Atmospheric correction methodologies that use radiative transfer algorithms and atmospheric characterization data are required for automated large area processing. For combined sensor use a consistent radiative transfer algorithm and atmospheric characterization for both sensors is required. Recently, a new algorithm, now named LaSRC (Land Surface Reflectance Code), was developed (Vermote et al. 2016) and has been adapted for Sentinel-2 application. It uses the 6S radiative transfer model, but has improved aerosol determination, using a ratio between the red and blue bands rather than between the middle infrared and blue bands and a linear function of a spectral index with slope and interpret terms derived from a spatially explicit climatology of MODIS and MISR data. The LaSRC has higher accuracy than LEDAPS for Landsat-8 application (Vermote et al. 2016) and performed well for Landsat-8 and Sentinel-2A application in the 11-12th April 2017 CEOS-WGCV Atmospheric Correction Inter-comparison Exercise, Frascati, Italy.

We obtained operational LaSRC source code (V3.5.5) and applied it with encouraging results to Landsat-8 and Sentinel-2A. We used it to characterize and quantify spectral differences between the Landsat-8 and Sentinel-2A bands for TOA, surface reflectance, and nadir BRDF-adjusted surface reflectance (NBAR). The results were published in a Sentinel-2 special edition of Remote Sensing of Environment (Zhang et al. 2018).

3.5 Cloud and water masking

Cloud masking is needed to discard cloud contaminated observations. The Landsat-8 Collection 1 cloud mask, that improves on the previous Landsat-8 cloud mask, was used. Landsat-8 30 m pixels that were labelled as cloudy with high or medium confidence were discarded. We found that the current Sentinel-2A L1C cloud product performs poorly, and so the latest SEN2COR (V2.3.1) 20 m cloud mask was applied independently to each L1C tile acquisition using the default parameter settings. The SEN2COR cloud mask was found to be reliable over Southern Africa in another recent study (Roy et al. 2017). The 20 m Sentinel-2A pixels (or four 10 m pixels) labelled as high or medium confidence cloud were discarded. No masking for shadows was implemented as they are difficult to detect reliably and are often confused with burned areas.

Water bodies can be confused with burned areas in single date images and so were detected and discarded from the analysis but in a conservative way to avoid removal of burned pixel observations. Water has low and monotonically decreasing reflectance with wavelength, particularly in the 550 nm to 900 nm wavelength range. In general, SWIR reflectance is lower than green reflectance over water, in contrast to other dark surfaces such as burned areas (Huang et al. 2016). Thus, water pixels were selected as those that had \( \rho_{550} < \rho_{665} \) and \( \rho_{550} > \rho_{1640} \). We found that application of the \( \rho_{550} > \rho_{665} > \rho_{1640} \) test provided more conservative water masking when applied to the TOA reflectance rather than to surface reflectance. This is because over dark surfaces the shorter wavelength bands are particularly
sensitive to aerosol effects and are less reliably atmospherically corrected than at longer wavelengths. Examples of the water masking are shown in Figures 6 and 7.

3.6 Reflectance correction to nadir BRDF-adjusted reflectance (NBAR)

The majority of terrestrial surfaces reflect optical wavelength radiation with a directional dependence that varies as a function of the sun–target–sensor geometry, commonly described by the Bi-directional Reflectance Distribution Function (BRDF). We investigated and quantified view zenith BRDF effects in Landsat (Roy et al. 2016) and in Sentinel-2A (Roy et al. 2017a, b). Across the Landsat swath, the red and NIR reflectance can vary by up to 0.02 and 0.06 (reflectance units) due only to BRDF effects, and up to 0.06 and 0.08 for Sentinel-2 because of its wider field of view. These differences may constitute a significant source of noise for certain applications including classification and change detection. We developed code to correct for this using a semi-physical approach that we developed for Landsat application (Roy et al. 2016) and demonstrated also works for Sentinel-2A (Roy et al. 2017a,b). Rather than use empirical solutions that require the presence of similar land cover types or pseudo-invariant features located across each image that are often not available, the gridded surface reflectance is adjusted to a nadir view for a specified solar zenith to provide nadir BRDF-adjusted reflectance (NBAR). A $c$-factor approach is used based on multiplying the surface reflectance with the ratio of the reflectances modeled using MODIS BRDF spectral model parameters for the observed and for a nadir view and a specified solar zenith. Our research demonstrated that the BRDF shapes of different terrestrial surfaces are sufficiently similar over the Landsat or Sentinel-2 field of view that a fixed set of global average MODIS BRDF spectral model parameters is adequate. The adjustment is conservative with low sensitivity to the land cover type, condition, or surface disturbance, which is important given the subsequent time series application and because the BRDF of vegetated surfaces changes post-fire.

3.7 Correction for non-random sensor differences

The above processing introduces small non-random differences when the surface NBAR Sentinel-2A MSI and Landsat-8 OLI data are compared (Zhang et al. 2018). Consequently, the Sentinel-2A MSI surface NBAR were adjusted to be more consistent with the Landsat-8 OLI surface NBAR. The band-specific NBAR linear regression parameters described in Zhang et al. (2018) were used. The parameters were derived by ordinary least squares (OLS) regression of >65 million Sentinel-2A and Landsat-8 samples extracted with per-pixel quality filtering over the southern Africa study area considered in this study and processed to surface NBAR as described above. The OLS regressions were significant ($r^2 > 0.9$ and $p < 0.0001$) for the bands used in this study and in general the OLI surface NBAR is lower than the MSI surface NBAR, by a factor of 0.94 (red band) to 0.99 (NIR band), except for the green band that is almost similar between sensors (Zhang et al. 2018). The sources sensor surface NBAR differences are complex and include residual atmospheric correction errors in the LaSRC atmospheric correction and in the subsequent the adjustment to NBAR process. In addition, although both sensors are well calibrated, non-random residual calibration errors remain when the sensor data are compared (Helder et al. 2018). We note also that the spectral band passes of the equivalent Sentinel-2A MSI and Landsat-8 OLI bands are quite similar but are not identical (Zhang et al. 2018). Consequently, reflectance
retrieved from Sentinel-2A and Landsat-8 sensed under the same conditions will be different to a degree depending on the spectral variation of the reflectance with respect to the MSI and OLI spectral response functions.

### 3.8 Preparation for Sentinel-2B availability

Sentinel-2A was launched on June 23rd 2015 and Sentinel-2B was launched March 7th 2017. Both are in circular sun-synchronous 786 km orbits with 98.62° inclination and equatorial crossing times of 10:30AM and with a phase delay of 180°. A study was undertaken to examine the improved temporal resolution provided by integration of the three sensors (Li and Roy 2017). The study involved a global annual analysis of Landsat-8, Sentinel-2A and Sentinel-2B metadata obtained from the committee on Earth Observation Satellite (CEOS) Visualization Environment (COVE) tool. A global equal area projection grid defined every 0.05° was used considering each sensor and combined together. Histograms, maps and global summary statistics, of the temporal revisit intervals (minimum, mean, and maximum) and of the number of observations were reported. Figure 5 illustrates the average revisit interval for 2016 considering the different sensor combinations. Table 1 summarizes the global mean, median, and the first and second most frequent average revisit intervals. When each sensor is considered alone the globally most frequent average revisit intervals are the same as the median values and correspond to the sensor repeat cycles, i.e., 16 days (Landsat-8) and 10 days (Sentinel-2A or -2B), and occur for 54.6% (Landsat-8) and about 56% (either Sentinel-2) of the globe.

<table>
<thead>
<tr>
<th>Sensor Combination</th>
<th>Mean (decimal days)</th>
<th>Median (decimal days)</th>
<th>Most Frequent Total Value (%)</th>
<th>2nd Most Frequent Total Value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat-8</td>
<td>12.130</td>
<td>16.000</td>
<td>16.000 (54.6%)</td>
<td>7.972 (10.7%)</td>
</tr>
<tr>
<td>Sentinel-2A</td>
<td>7.771</td>
<td>10.000</td>
<td>10.000 (55.5%)</td>
<td>5.000 (14.0%)</td>
</tr>
<tr>
<td>Sentinel-2B</td>
<td>7.820</td>
<td>10.000</td>
<td>10.000 (56.6%)</td>
<td>5.000 (15.1%)</td>
</tr>
<tr>
<td>Landsat-8 + Sentinel-2A</td>
<td>4.593</td>
<td>4.611</td>
<td>6.097 (14.8%)</td>
<td>3.055 (1.8%)</td>
</tr>
<tr>
<td>Landsat-8 + Sentinel-2B</td>
<td>4.593</td>
<td>4.611</td>
<td>6.097 (14.9%)</td>
<td>3.055 (2.1%)</td>
</tr>
<tr>
<td>Sentinel-2A + Sentinel-2B</td>
<td>3.795</td>
<td>3.667</td>
<td>5.000 (29.0%)</td>
<td>3.333 (13.6%)</td>
</tr>
<tr>
<td>Sentinel-2A + Sentinel-2B</td>
<td>3.795</td>
<td>3.667</td>
<td>5.000 (29.0%)</td>
<td>3.333 (13.6%)</td>
</tr>
<tr>
<td>Sentinel-2B + Sentinel-2B</td>
<td>2.835</td>
<td>2.858</td>
<td>3.792 (11.8%)</td>
<td>2.750 (3.1%)</td>
</tr>
</tbody>
</table>

Table 1. Global summary statistics (mean and median) of the average satellite revisit interval from January 1st to December 31st 2016 for different satellite combinations reported in decimal days (to 3 d.p.). In addition, the first, and second most frequent average satellite revisit interval values are tabulated with the percentage of global grid points with that value in parenthesis. Results derived from the global 0.05° data (Fig. 5).
Figure 5. The average satellite revisit interval (days) from January 1st to December 31st 2016 for (a) Sentinel-2A, (b) Landsat-8, (c) Landsat-8 and Sentinel-2A, (d) Sentinel-2A and Sentinel-2B, (e) Landsat-8, Sentinel-2A and Sentinel-2B. Global results derived at 7201 × 3601 points spaced every 5.559752 km, equivalent to 0.05° at the Equator, in the equal area sinusoidal projection.

When the sensors are combined the overlap of their different orbit swaths decreases the average revisit intervals. The globally most frequent average revisit intervals are 6.1 days (Landsat-8 and either Sentinel-2 sensor), 5.0 days (both Sentinel-2 sensors), and 3.8 days (all three sensors), and occur for about 14.8%, 29.0% and 11.8% of the globe respectfully. Landsat-8 and either Sentinel-2 sensor together have global mean and median average revisit intervals of a 4.6 and 4.5 days respectively; the two Sentinel-2 sensors together have global mean and median average revisit intervals of 3.8 and 3.7 days respectively; and for all three sensors the global mean and global median average revisit interval is 2.9 days.

The temporal observation frequency improvements afforded by sensor combination are significant. For example, if we conservatively assume that 50% of observations are cloudy at the time of satellite overpass than we can expect on average to have a global cloud-free observation better than every 6 days when Landsat-8, Sentinel-2A and -2B are combined. This will enable meaningful Landsat-8 Sentinel-2 time series based burned area detection.
4.0 Development of an automated sensor-agnostic burned area algorithm

4.1. Algorithm
The developed burned area mapping algorithm is based on a multi-temporal change detection approach and uses every available Landsat-8 and Sentinel-2 acquisition. Seamless integration of the different sensor data, and the burned area mapping, is achieved through a random forest change regression, parameterized with synthetic training data and modeling reflective wavelength change due to fire. The algorithm is applied on a per-pixel basis to the gridded 30 m surface NBAR. It can be run with any temporal combination of Landsat-8 and Sentinel-2 observations. This is necessary because the temporal availability of cloud-free Landsat-8 and Sentinel-2 data may be quite variable. As residual cloud detection errors (omission and commission errors) and atmospheric correction errors (over or under correction) will occur the algorithm is designed to be robust to these issues and also to the presence of shadows that cannot be detected reliably. Temporal consistency checks are used to reduce burned area mapping commission errors and a region growing algorithm is used to reduce omission errors.

Burned area mapping is achieved using a random forest regression estimator parameterized with synthetic training data that models reflective wavelength change due to fire. A burn candidate can only be detected if there are at least three observations, spaced no more than 15 days apart, over a 30 day period prior to and after fire occurrence. A conservative region growing algorithm is subsequently used to reduce burn mapping omission errors due to temporally sparse observations. In the resulting burned area product, each 30 m gridded pixel location is classified as burned, unburned, or unmapped. At the burned pixel locations the estimated day of burning is also defined, together with an estimate of the product of the subpixel fraction burned \((f)\) and the combustion completeness \((cc)\). The algorithm is described in detail in Roy et al. (2018).

Fig. 6 illustrates a detailed example of the algorithm functioning, showing gridded 30 m Sentinel-2A and Landsat-8 surface 30 m surface NBAR that were sensed one day apart, and the results of the random forest regression applied to the two dates of data. The satellite data were processed as described in Section 3. In this example, pixels that were unmapped are quite evident (Fig. 6c, grey) and occur along the Kafue River where the Sentinel-2A and/or Landsat-8 observations were flagged as water. As expected, only the regions that burned between the two dates have high \(f,cc\) values, and in general the values are often smaller at the edges of the burn, where the sub-pixel fraction burned \((f)\) is small.
Figure 6. False color 30 m surface NBAR (~2.2 µm, 1.6 µm, 0.86 µm) (a) Sentinel-2A June 2nd 2016, (b) Landsat-8 June 3rd 2016; and the f.cc (c) estimated using the random forest regression applied independently at each 30 m pixel location to the two dates of data. Results shown for 2000 × 2000 30 m pixels (60 × 60 km) in the north east corner of Kafue National park, Zambia (centered on 14.6634°S, 26.1352°E).

Fig. 7 shows the random forest regression results as Fig. 6 but with respect to Sentinel-2A surface NBAR data sensed 10 days, rather than one day, later. The progression of burning and a general drying of the landscape (higher unburned surface NBAR in the later image) are apparent when Figs. 6 and 7 are compared. The older burns have less evident burn signatures. Thus, for example, the f.cc values over the central burned area in Fig. 7, are generally smaller in Fig. 6 because the burn signal was reduced 10 days after the fire occurred compared to one day after.
Figure 7 False color 30 m surface NBAR (~2.2 µm, 1.6 µm, 0.86 µm) (a) Sentinel-2A June 2nd 2016, (b) Sentinel-2A June 12th 2016; and the f.cc (c) estimated using the random forest regression applied independently at each 30 m pixel location to the two dates of data. Results shown for 2000 × 2000 30 m pixels (60 × 60 km) in the north east corner of Kafue National park, Zambia (centered on 14.6634°S, 26.1352°E).

The 30m burned area product is generated on a monthly basis for reporting convenience, and to allow easy comparison with the monthly NASA MODIS burned area product. For example, six months of Landsat-8 and Sentinel-2A data, are used to generate four monthly products (Figure 8).
At the burned pixel locations, the estimated day of burning in the month is defined with an estimate of the $f_{cc}$ (except for the region grown burned pixels where the $f_{cc}$ is undefined).

![Figure 8](image)

**Figure 8** Illustration of the data time periods used to generate monthly burned area estimates.

4.2. Example results and comparison with the MODIS 500m burned area product

Fig. 9 shows the area used to prototype the burned area mapping algorithm. It covers approximately $1112 \times 1112$ km of southern Africa and is bounded by latitudes $10^\circ$S to $20^\circ$S and longitudes $20.31^\circ$E to $31.93^\circ$E; it corresponds to the area defined by MODIS land tile h20v10 (Wolfe et al. 1998). It also shows an orbit of Landsat-8 and Sentinel-2A data. Fewer observations are provided by Landsat-8 compared with Sentinel-2A due to the relatively smaller swath width and lower repeat cycle of Landsat-8. The benefit of using both sensors together to provide a richer time series is evident. In July 2016, the median number of observations that were not flagged as cloud or as water was 2 (Landsat-8), 3 (Sentinel-2A), and 5 (both Sentinel-2A and Landsat-8).

This area illustrated in Figure 9 is fire prone, and is well known to the PI who participated in the NASA SAFARI 2000 campaign that was held in this region. The region is characterized by extensive annual burning, and includes large and small burned areas that are often not temporally persistent and so are challenging to map reliably. Most fires occur in the dry season from approximately June to September when herbaceous vegetation is either dead (annual grasslands) or dormant, and when deciduous trees have shed their leaves, contributing to an accumulation of dry and fine fuels that are easily combustible. The majority of the fires are anthropogenic, lit for numerous reasons including maintaining pasture and clearing land, with a minority of lightning ignited fires associated with early wet season thunderstorms.
Figure 9. Study area over southern Africa corresponding to MODIS land tile (h20v10) covering 37,065 × 37,065 30 m pixels. The grey lines illustrate the boundaries of the 49 sinusoidal projection 5295 × 5295 30 m pixel tile boundaries. The white lines show national boundaries. False color (~2.2 µm, 1.6 µm, 0.86 µm) surface nadir BRDF-adjusted reflectance (NBAR) for a single Landsat-8 OLI orbit (185 km wide, left) and a single Sentinel-2A MSI orbit (290 km wide, right) that were both sensed July 19th 2016 are shown.
Fig. 10 shows the mapped burned areas for July 2016 for the northwest 5295 × 5295 30 m pixel tile (top left in Fig. 9) that falls in Lunda Sul Province, Angola. The left column shows the results derived using Sentinel-2A only and the right column shows the results derived using both Sentinel-2A and Landsat-8 surface NBAR time series. Results for Landsat-8 alone are not shown because usually there are only two Landsat-8 observations per month which often precludes Landsat-8 only burned area mapping. In Fig. 10, the approximate day of burning (top row) is shown with a rainbow coloring scale, from the beginning (blue tones) to the end (red) of July. The mapped days of burning exhibit a coherent spatio-temporal pattern. Notably, the same patterns are observed for the Sentinel-2A only (left) and the combined Sentinel-2A and Landsat-8 (right) results but with more pixels labelled as burned and sometimes a different estimated day of burning in the combined sensor results. A total of 17.6% and 21.0% of the pixels were labelled as burned by the Sentinel-2A and combined sensor results respectively. There were more unmapped pixels in the Sentinel-2A (0.03%) than in the combined sensor (0.01%) results, but because there are so few and they are scattered, they are not apparent in the figure. The $f_{cc}$ values (middle row) are shown for $f_{cc} \geq 0.2$ (i.e., the minimum $f_{cc}$ threshold value) and exhibit a range of $f_{cc}$ from 0.2 to 0.96. The Sentinel-2A and the combined sensor $f_{cc}$ values are similar where burning was detected. The region growing results (bottom row) are also similar but with more Sentinel-2A region-grown pixels (0.55%) compared to the combined sensor results (0.16%). All of these differences reflect the greater temporal density of observations provided by combined Sentinel-2A and Landsat-8 use and indicate that the Sentinel-2A and Landsat-8 sensor data are pre-processed and integrated into the mapping algorithm in a coherent manner.

Throughout the burning season, the physical characteristics of burned areas in southern Africa change as the fuel condition and ambient conditions (e.g., wind speed and humidity) change. This is evident in Fig. 11 that shows the mapped burned areas for July, August, September. The day of burning defined by the Collection 6 MODIS 500 MCD64A1 burned areas product (left column) and the Sentinel-2A and Landsat-8 results (right column) are shown. The MODIS burned area product is produced using a hybrid algorithm that combines daily Terra and Aqua MODIS surface reflectance imagery with daytime and nighttime MODIS active fire detections to map burning to the nearest day within 500 m pixels on a monthly basis (Giglio et al. 2018). The MODIS product was resampled from 500 m to 30 m by nearest neighbor resampling to preserve the pixel values. In both products, there are slightly more burned areas mapped in August than July, and the least in September when the fire season starts to finish. The products exhibit a similar spatio-temporal mapped day of burning for each month. The MODIS product uses daily satellite data and so is likely to more precisely define the day of burning.
Figure 10 Burned area 30 m mapping results for July 2016: (top row) the day of burning colored as days after the first day of the month: 0-2, 3-5, 6-8, 9-11, 12-14, 15-17, 18-20, 21-23, 24-27, 28-31; (middle row) f.cc colored as Figs. 6c and 7c; (bottom row) region growing results colored as: unburned (black), burned pixel (magenta), seed burned pixel (sea green), region grown burned pixel (light blue). The left column shows the results derived using Sentinel-2A surface NBAR only. The right column shows the results derived using both Sentinel-2A and Landsat-8 surface NBAR. Results shown for the 5295 × 5295 30 m pixels (158.85 × 158.85 km) defining the northwest study area tile (Fig. 9) located over Lunda Sul Province, Angola.
Figure 11 Comparison of the MODIS MCD64A1 burned area product day of burning (left column) and the Sentinel-2A and Landsat-8 day of burning (right column) in 2016 for July (top row), August (middle row) and September (bottom row), colored as the top row of Figure 10 for each month. Results shown for the 5295 × 5295 30 m pixels defining the northwest study area tile (Fig. 9) located over Lunda Sul Province, Angola.
Figures 12 and 13 show the MODIS MCD64A1 500 m and the Sentinel-2A and Landsat-8 30 m burned area results respectively for July 2016. The MODIS burned area product was resampled from 500 m to 30 m by nearest neighbor resampling to preserve the pixel values. The spatio-temporal distribution of burning is similar between the two products. Areas of extensive burning in the center (Zambia), the northwest (Angola), and in and around the Caprivi Strip (Namibia) are apparent in both products. Areas with no burning, either where the fire season has not yet started in the southern parts of the study area, or where there was insufficient fuel due to removal by humans and animals or due to arid conditions, are evident particularly in Zimbabwe and Botswana. Unmapped areas over the water bodies and wetlands are evident in both products. Even at this scale the greater number of small burned areas in the Sentinel-2 and Landsat-8 product compared to the MODIS product is apparent.
Figure 12  July 2016 day of burning detected by MODIS MCD64A1 Collection 6 500 m burned area product for the study area (Fig. 9), colored as days after the first day of the month: 0-2, 3-5, 6-8, 9-11, 12-14, 15-17, 18-20, 21-23, 24-27, 28-31. The grey tones show pixels that were not mapped by the MODIS algorithm (Giglio et al. 2018) and occur predominantly over water bodies and wetlands, notably, the Okavango Delta (Botswana), Lake Kariba and Lake Bangweulu (Zambia), and Lake Cahora Bassa (Mozambique). The figure was generalized from 30 m to 450 m resolution by taking the majority value in 15 × 15 30 m pixel windows.
Figure 13 July 2016 day of burning detected considering Sentinel-2A and Landsat-8 surface NBAR time series for the study area (Fig. 9), colored as days after the first day of the month: 0-2, 3-5, 6-8, 9-11, 12-14, 15-17, 18-20, 21-23, 24-27, 28-31. The grey tones show pixels that were unmapped. The figure was generalized from 30 m to 450 m resolution by taking the majority value in $15 \times 15$ 30 m pixel windows.

Table 2 shows a confusion matrix and accuracy metrics that compare the July MODIS MCD64A1 burned area product (Fig. 12) with the Landsat-8 and Sentinel-2A (Fig. 13) burned area results. The MODIS MCD64A1 product maps 65,170 km$^2$ of burning across the study area, whereas the Landsat-8 and Sentinel-2A product maps 112,454 km$^2$, resulting in a -41.9% relative bias. The omission and commission errors are 0.64 and 0.39 respectively, reflecting the presence of small burned areas in the Sentinel-2A and Landsat-8 product that are missing in the MODIS product, as apparent in Figs. 12 and 13 (and in Fig. 11). Less than 0.9% of the study area was unmapped by
either product. The overall accuracy (i.e., the overall proportion of pixels correctly classified either as burned or unburned) is high (0.92) and is typical of binary burned area classifications and reflects the prevalence of the unburned class.

Table 2 Burned area mapping confusion matrix for July 2016 comparing Landsat-8 and Sentinel-2A and 30 m mapping results with Collection 6 MODIS MCD64A1 500 m results resampled to 30 m. The two burned area maps are compared over the whole study area (Figs. 12 and 13), covering 37065 × 37065 30 m pixels (1112.55 km × 1112.55 km). For the computation of the accuracy metrics, the Landsat-8 and Sentinel-2A results are treated as reference data, and the MCD64A1 results are treated as classified data.

<table>
<thead>
<tr>
<th>Landsat-8/ Sentinel-2A</th>
<th>MCD64A1</th>
<th>Accuracy Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burned [km²]</td>
<td>Burned</td>
<td>Overall Accuracy (0-1)</td>
</tr>
<tr>
<td>(3.22%)</td>
<td>39871.5</td>
<td>Omission Error (0-1)</td>
</tr>
<tr>
<td>Unburned [km²]</td>
<td>Unburned</td>
<td>Commission Error (0-1)</td>
</tr>
<tr>
<td>(5.86%)</td>
<td>72427.8</td>
<td>User’s Accuracy (0-1)</td>
</tr>
<tr>
<td>Unmapped [km²]</td>
<td>Unmapped</td>
<td>Producer’s Accuracy (0-1)</td>
</tr>
<tr>
<td>(&lt;0.1%)</td>
<td>154.5</td>
<td>Relative Bias [%]</td>
</tr>
<tr>
<td>Row total [km²]</td>
<td>Row total</td>
<td></td>
</tr>
<tr>
<td>112453.9</td>
<td>1113972.5</td>
<td></td>
</tr>
</tbody>
</table>

The confusion matrix results reported in Table 2 may contain biases due only to the different 500 m (MODIS) and the 30 m (Landsat-8 and Sentinel-2A) reporting resolutions. For this reason, Fig. 14 shows for all the study area, a scatterplot of the proportions of grid cells labeled as burned by the MODIS burned area product plotted against the proportions labeled as burned by the Sentinel-2A and Landsat-8 product. Geographical analyses of this kind are sensitive to the size of the grid cells, using increasing larger grid cell sizes will result in higher statistical measures of correlation between the products. Grid cells with side dimensions of 3 km, 4 km, 5 km, and 6 km were used as they are many times larger than the 500 m MODIS pixel size while being sufficiently small to reduce the occurrence of cells with similar burned-area proportions in both products but with burns mapped at different locations within the cell. Only grid cells with < 50% unmapped data were considered. A rainbow logarithmic color scale is used to illustrate the frequency of cells having the same x- and y-axis proportion values. This is an established comparison approach - the slope and offset of the regression line fitted through the data are indicative of the accuracy of the burned area detection, and the coefficient of determination ($r^2$) is indicative of the precision (Giglio et al. 2018). As for the confusion matrices, the Landsat-8 and Sentinel-2A mapping results are treated as reference data and the MODIS MCD64A1 results are treated as classified data.

The correlation between the MODIS MCD64A1 and the Landsat-8 and Sentinel-2A cell proportions increases monotonically with the cell size, from $r^2 = 0.58$ (3 km cells) to $r^2 = 0.65$ (6 km cells), reflecting how the omission and commission errors compensate each other in the progressively coarser cells. The slopes of the regression lines are very similar, varying between 0.750 and 0.762, and together with the negative intercept, indicate that for all four cell sizes, the MODIS product maps a smaller proportion of the landscape as burned than the Landsat-8 and Sentinel-2A product. The small size and high fragmentation of the burned areas mapped by both products is reflected by the density plots: for all four cell sizes, the majority of points (colors from
green to red) correspond to low burning proportions, and noticeably at the coarser 6 km resolution only a few 6 km cells are fully burned.

**Figure 14** Scatter plots of the proportions of grid cells labeled as burned by the MODIS MCD64A1 Collection 6 burned area product (Fig. 12), plotted against the proportion labeled as burned by the Sentinel-2A and Landsat-8 30 m product (Fig. 13), for the study area in July 2016. The point density, calculated using a 30 × 30 quantization of the plot axes, is displayed with a rainbow logarithmic color scale. Results are shown considering grid cells with side dimensions of (a) 3 km, (b) 4 km, (c) 5 km, and (d) 6 km. The blue lines show the ordinary least squares regression of the plotted data. The dashed 1:1 lines are shown for reference.
Figure 15 shows a histogram of the temporal difference in the day of burning in July 2016 mapped by the 500 m MODIS burned area product (nearest neighbor resampled to 30 m) and the Sentinel-2 and Landsat-8 product where both detected a burned area at 30 m. As noted earlier, the MODIS burned area product uses daily satellite data and so will more precisely define the day of burning. The histogram is approximately normally distributed with a mean and median difference of 3.37 and 3.00 days respectively. Evidently, the Sentinel-2A and Landsat-8 product reports the day of burning on average about three days later than the MODIS product. This later reporting is expected because of the reduced revisit of the Sentinel-2A and Landsat-8 compared to MODIS, and because the Sentinel-2A and Landsat-8 algorithm reports the day of burning as having occurred on the first date of NBAR observation after the change due to fire.

Figure 16 shows the same scatterplot as Fig. 14 but was derived defining the Sentinel-2A and Landsat-8 July mapped burned areas as those that were mapped three days later, i.e., from July 4th to August 3rd 2016, and not from July 1st to 31st 2016. The regression slopes and $r^2$ values are all increased. This is because of reduced errors of omission and commission in the reporting of the burning that occurred at the beginning and end of the month. This is evident in the reduced number of grid cells clustered along the vertical and horizontal plot axes that correspond to small relative proportions of the grid cells mapped as burned by Sentinel-2A and Landsat-8 compared to the MODIS MCD64A1 product, and vice versa.
Figure 16 As Fig. 14 but with a three-day adjustment applied to the Sentinel-2A and Landsat-8 30 m results according to the date of burning difference shown in Fig. 15.

6.0 Future work

6.1 Development of 30 m burned area product validation

The above comparisons, and the figures comparing the products, provide confidence that the burned area mapping algorithm is producing coherent results. This is not the same, however, as validation that requires the use of independent reference data collected with no, or minimal error, and with a suitable spatio-temporal sampling design to derive statistically rigorous accuracy measures.

We negotiated free-access to 3 m commercial PLANET (NIR, red, green, blue wavelength) images over Africa and have begun to order two date image pairs and we have developed code to reproject the PLANET images into registration with the tiles. Figure 17 illustrates a preliminary validation result for a 25.8 km × 15 km area. The two PLANET acquisitions were sensed 31 days
apart and captured a significant number of burns that occurred within that period. The 1.8 km × 1.8 km subsets (i.e., 600 × 600 3m PLANET pixels, and 60 × 60 30 m pixels) illustrate the complexity of the landscape and how very small and fragmented burned areas are present. The spatial pattern of the 30 m $f_{cc}$ generally coincide with the burns evident in the two-date PLANET data.

The variation of the 30 m $f_{cc}$ evident in Fig. 17 also indicates that the algorithm may capture the unevenly spatially distributed combustion completeness ($cc$). Validation of combustion completeness requires ground based assessment. We demonstrated the rapid assessment of grass biomass (Cooper et al. 2017) as step in this direction using structure-from-Motion (SfM) photogrammetry and Terrestrial Laser Scanning (TLS) and showed that these approaches were rapid and more accurate than conventional disc pasture meter and allometry approaches.

These preliminary validation results in a complex environment are encouraging. The ability to detect small burns is particularly needed to help resolve scientific debate concerning the contribution of small fires to pyrogenic emissions.

Future product accuracy assessment will be conducted following the validation protocol endorsed by the Committee on Earth Observation Satellites (CEOS). Burned area maps interpreted from high spatial resolution commercial satellite two date image pairs will be compared with the 30 m burned area product to derive quantitative accuracy metrics. Given commercial data availability issues only a CEOS Stage 2 validation, i.e., “Product accuracy has been assessed over a widely distributed set of locations and time periods, representative of the full range of conditions present in the product” is feasible.
Figure 17. Zambia validation results for a 25.8 km × 15 km area in Zambia. (1) PLANET 3 m pixel images acquired July 18th (day 200) 0.82/0.63/0.54 μm false color – burns appear blue/black, (2) PLANET image acquired August 18 (day 231), (3) the Lansat-8/Sentinel-2A 30 m f.cc colored as Fig. 6c and 7c, (4) 30 m day of burning (colored: 200≤doy ≤204, 210≤doy ≤214, 215≤days ≤219, 220≤doy ≤224, 225≤days ≤229, 230≤doy ≤234). (a), (b), (c) and (d) show very detailed 1.8 km × 1.8 km subsets (a and b are 3 m, c and d are 30 m).
6.2 Investigation of 30 m burned classified pixel \( f.cc \) retrievals

In the resulting burned area product, each 30 m pixel is labeled as burned, unburned or unmapped, with an estimate of the day of burning and the \( f.cc \) at the burned pixel locations. We note that the \( f.cc \) data provide a burned area characterization beyond the conventional binary burned/unburned classification. This may translate into more accurate estimates of biomass burned, needed for estimating pyrogenic emissions using bottom-up approaches, and may complement satellite derived fire severity measures used by the post-fire assessment and remediation applications community. Future research to investigate this is recommended.

6.3 Scaling to Africa wide production

The burned area mapping algorithm is automated as even modest levels of user intervention are not scalable to the considerable data volumes provided by the Landsat-8 and Sentinel-2 sensors. A new proposal “Africa burned area product generation, quality assessment and validation – demonstrating a Multi-Source Land Imaging (MuSLI) Landsat-8 Sentinel-2 capability” has been awarded (NNH17ZDA001N-LCLUC Land-Cover/Land-Use Change). Monthly 30 m burned area products for 2017, 2018, and 2019 will be generated using Landsat-8 and Sentinel-2A and also using Sentinel-2B for 2018 and 2019 (Table 3). This schedule reflects the availability of Sentinel-2A data in 2017+ and the later availability of Sentinel-2B data in 2018. The Table 3 data volumes were derived as the product of the number of files (Fig. 18) and the mean 964.69 MB (Landsat-8 Collection 1) and 471.86 MB (Sentinel-2A L1C) compressed file sizes considering all the Africa 2016 data in our archive. The total input volume is 298 TB. Currently we are investigating the use of the NASA harmonized Landsat Sentinel-2 (HLS) products as an input, and processing on Amazon Web Services.

![Figure 18 Burned area product generation for all of Africa south of 23.4 N (red line, the Tropic of cancer) encompassing:](image)

**Left:** 1041 Landsat-8 WRS-2 path/rows,

**Right:** 2829 Sentinel-2 L1C tiles.

| Table 3: Study area annual input number of Landsat-8 and Sentinel-2 files (data volume in parentheses) |
|-------------------------------------------------|-----------|-----------|-----------|----------|
| Landsat-8 Collection 1                          | 23747 (21.8 TB) | 23747 (21.8 TB) | 23747 (21.8 TB) | 71241 (65.5 TB) |
| Sentinel-2A L1C                                 | 103258 (46.5 TB) | 103258 (46.5 TB) | 103258 (46.5 TB) | 309774 (139.5 TB) |
| Sentinel-2B L1C                                 | -          | 103258 (46.5 TB) | 103258 (46.5 TB) | 206516 (93 TB) |
7.0 Publications and presentations

The following articles were published:


The following articles were submitted:


The following presentations were made:


3. Roy, D.P., Huang, H., Boschetti, L., Zhang, H., Yan, L., Li, Z., Landsat-8 Sentinel-2 global burned area product Prototyping (Type II) to Production (Type 1), *NASA Land Cover Land Use*
Change Spring Science Team Meeting, Gaithersburg Marriott Washingtonian Center (Rio), Maryland, April 3-5, 2018.


