Validation of MCD64A1 and FireCCI51 cropland burned area mapping in Ukraine

Joanne V. Hall *, Fernanda Argueta, Louis Giglio

University of Maryland, Department of Geographical Sciences, 2181 LeFrak Hall, College Park, MD 20742, USA

**ARTICLE INFO**

**ABSTRACT**

Small fires represent an important under-represented fire type within active fire and burned area datasets in addition to fire emission inventories, especially in regions with substantial agricultural areas. In order to improve regional and global burned area and fire emissions inventories, active fire and burned area algorithm developers are focusing on improving the mapping accuracy of the timing and spatial extent of these small fires. However, product developers have often relied on burned area validation methods that are designed for larger wildfires, and which are therefore not appropriate for these small fire types. Specifically, validation of crop residue burns using pre- and post-Landsat and Sentinel imagery (as recommended by the Committee on Earth Observing Satellites Working Group on Calibration and Validation (CEOS) Land Product Validation protocol) breaks down because the short duration (less than 1 day) between the burn and the subsequent field plowing removes the burned signature before the next sensor overpass. Here we describe an alternate approach that allowed us to rigorously validate two widely available, coarse-resolution global burned area products — MCD64A1 and FireCCI51 — in cropland through using exhaustively-mapped field-level burned area reference maps produced for seven reference areas in Ukraine in 2016 and 2017. Our results highlight the overall high omission errors (MCD64A1: 71–76% and FireCCI51: 63–99%) and commission errors (MCD64A1: 62–81% and FireCCI51: 49–93%) for both products within cropland, while also demonstrating the difficulty of mapping crop residue burned area within the spring, pre-planting mapping period compared to the summer, post-harvest mapping period. Product-specific artifacts and errors are also demonstrated including the confusion with the larger harvest spectral signal (MCD64A1) and the large swaths of unmapped pixels clustered in regular geometric shapes (FireCCI51). These validation results will be used to help guide the upcoming MCD64A1 Collection 7 burned area product improvements within cropland.

1. Introduction

The importance of accurately representing small fires within burned area (BA) and active fire (AF) datasets has been garnering increased attention within the fire-science community. Over the past several years, attempts have been made with varying success to improve the representation of small fires within global BA datasets (e.g., Randerson et al., 2012; Ramo et al., 2021; Roteta et al., 2019; van der Werf et al., 2017). Small fires represent an important under-represented source within fire emission inventories, especially in regions with substantial agricultural areas. Although small fires emit relatively low quantities of emissions, collectively they have impacts that far exceed the fire location (e.g., Hall and Loboda, 2017; Zhou et al., 2018). In particular, the proximity of inherently-small crop residue fires to urban locations, and the recurring nature of these managed fires, can negatively impact air quality and human health (e.g., Adler, 2010; Chakrabarti et al., 2019).

Accurately mapping the extent of small fires and subsequent emissions is a crucial next step in improving regional and global BA and fire emissions inventories. Although there are regional-based studies aimed at improving crop residue BA and emissions (e.g., Liu et al., 2020a; Hall et al., 2021), there has also been a move toward improving the small fire accuracies within global-based BA algorithms (e.g., Giglio et al., 2018) — a key input in several fire emission databases. According to the Global Fire Emissions Database with small fire boost (GFED4.1s), a large proportion of small fire BA is found in temperate and sub-tropical agricultural regions (van der Werf et al., 2017). Yet, BA validation within agricultural and cropland regions is typically not included in larger global studies (e.g., Boschetti et al., 2019) or is conducted with methods...
that are inappropriate for the short-lived, managed small fires that are characteristic of agricultural burning (e.g., Padilla et al., 2015).

Quantifying the accuracy of mapped BA, regardless of size, relies on extensive validation efforts designed with the specific fire type in mind. The current Committee on Earth Observing Satellites Working Group on Calibration and Validation (CEOS) Land Product Validation (LPV; https://lpvs.gsfc.nasa.gov/) protocol recommends using a higher resolution paired image approach (pre- and post-fire) to map the BA between the two acquisition dates (Boschetti et al., 2009). Although this method works very well for wildfires and larger prescribed burns (e.g., Boschetti et al., 2019; Padilla et al., 2015) it breaks down within cropland regions—a quintessential example of a small fire type.

Compared to the pixel size of moderate (20–30 m) and coarse (250 m – 1 km) resolution sensors, small fires only cover a fraction of the pixel area and often their optical burn signatures are not sufficiently distinct to be mapped with coarser resolution imagery. In addition, active fire products from polar-orbiting satellites are likely to miss such fires due to the limited diurnal sampling afforded by such platforms. Moreover, there are several limitations of mapping crop residue BA including 1) short duration fires with ensuing field plowing that removes the burned signature before the next sensor overpass (e.g., Hall et al., 2016, 2021), 2) partial field or piled residue burning that leads to less intense fires that are missed by coarse resolution sensors (e.g., Kulkarni et al. 2020; Lasko et al., 2017), and 3) increased BA fragmentation due to smaller landholdings within cropland regions (e.g., India - Liu et al., 2019; China - Zhang et al. 2018). These are only a few examples of why crop-residue BA products are prone to inaccuracies which lead developers of global BA products to accept higher small-fire omission errors rather than relax algorithm mapping thresholds and suffer increased commission errors throughout all land cover types (Roy and Boschetti, 2009). For example, the 500-m Moderate Resolution Imaging Spectroradiometer (MODIS) MCD64A1 Collection 6 burned area product (Giglio et al., 2018) improved the representation of small-fire BA overall, but at a cost of an increased commission error from confusion with the larger harvest spectral signal (Giglio et al., 2018; Zhang et al., 2018).

With the release of the European Space Agency Climate Change Initiative (CCI) Programme’s MODIS-based 250-m FireCCI51 BA product (Chuvieco et al., 2018; Lizundia-Loiola et al., 2020), which is targeted to capture small fires, there have been several studies that have compared and validated the MCD64A1 Collection 6 and FireCCI51 global BA products in various land cover types including those with frequent small (but not cropland) fires (e.g., Campagnolo et al., 2021; Fornacca et al., 2017; Humber et al., 2019; Shimabukuro et al., 2020; Valencia et al., 2020). A common theme throughout several of these papers is related to the differences in how these two hybrid BA algorithms utilize AF data to map BA pixels, and how these differences can lead to contrasting accuracy metrics. Overall, there seems to be an agreement that the improved algorithm in the MCD64A1 Collection 6 product outperforms the previous 5.1 generation for overall burned area mapping. However, the spectral confusion with harvest combined with the coarser resolution 500-m pixels reduces the mapping accuracy in cropland and other regions of predominantly small fires (e.g., Fornacca et al., 2017; Humber et al., 2019; Liu et al., 2019; Shimabukuro et al., 2020; Zhang et al., 2018; Zhu et al., 2017). In contrast, Campagnolo et al. (2021) found that the finer 250-m bands used for the FireCCI51 product helped improve the accuracy when detecting smaller fires (0–250 ha), but the lack of spectral range in those bands caused an increase in confusion errors for larger wildfires (>1000 ha). In addition, several studies have also noted the low temporal accuracy and bias towards the AF dates in the FireCCI51 product’s mapped BA date (e.g., Lizundia-Loiola et al., 2020).

With the steady publication of new BA products—some even designed for croplands (e.g., Wang et al., 2018)—and continual improvements within existing BA products, the lack of appropriate BA validation data and methods within cropland regions is both problematic and concerning. In particular, there are several countries with high concentrations of cropland area and long-standing histories of open-burning of crop residue for which remotely-sensed BA data are being consulted in policy-making decisions. For example, Ukraine consists of ~50% cropland area with very high levels of open-burning in March and April (pre-planting) and July - September (post-harvest; primary burning season) (Hall et al., 2021). Since the majority of open-burning in Ukraine is related to agricultural burning (e.g., crop residue, pastures, etc.) it is imperative to understand if this burning is captured accurately within BA products and fire emission databases.

We report here the first attempt to rigorously validate two widely available, coarse-resolution global burned area products—MCD64A1 and FireCCI51—in cropland using our novel database of exhaustively-mapped field-level burned area reference maps produced for seven reference areas in Ukraine in 2016 and 2017. The goal was to quantify the accuracy metrics of the two products within a densely-packed, large cropped field region to illustrate how well these MODIS-based BA products map small fires in an optimal cropland burning location given their coarse resolution.

1.1. Study area

Ukraine is one of the major global agricultural producers and exporters with ~50% of the land area used for arable production. The predominant soil type, Chernozem or “black soil”, is a fertile soil rich in humus (1.5–6% range; personal communication Dr. S. Maxim, National Scientific Center, Institute for Soil Science and Agrochemistry Research, Ukraine) and other essential plant nutrients which often leads to increased agricultural yields. This soil, along with the temperate continental climate, has allowed Ukraine to become one of the world’s top producers in sunflowers, corn/maize, wheat, sugar beet, barley, soy, and rapeseed. However, winter wheat (~17% total land area in 2017), sunflower (~14% total land area in 2017), and maize (~11% total land area in 2017) make up the majority of Ukraine’s cropland area (Kussul et al., 2017).

The winter wheat areas are located within the southern and eastern oblasts (i.e. administrative regions), the maize areas are located within the northern and western oblasts, while, sunflower crops are planted throughout Ukraine, often on the same fields as wheat or maize (Hall et al., 2021). Since farmers typically do not burn sunflower residue the two distinct burning peaks throughout the year are associated with preparing fallow fields for planting maize or sunflowers following the spring melt (March and April) or for removal of crop residue (i.e. winter wheat between July - September) before planting the next crop rotation.

2. Data and methods

2.1. MODIS burned area data

The MODIS burned area products we validated consist of 1) NASA’s MODIS Collection-6 MCD64A1 burned area product (Giglio et al., 2018) and 2) the FireCCI51 pixel-level burned area product developed within the European Space Agency’s CCI Fire Disturbance project (Chuvieco et al., 2018). Unlike the 500-m MCD64A1 product, the FireCCI51 product is generated at a spatial resolution of 250 m.

The MCD64A1 Burn Date layer was used to identify burned pixels (1–366), unburned pixels (0), unmapped pixels (~1), and water pixels (~2) during specified validation periods in 2016 and 2017 (Section 2.2). The FireCCI51 product uses a similar convention for its Julian Day of First Detection layer, from which we extracted the burn date (1–366) along with the unburned pixels (0), unmapped pixels (~1), and unburnable pixels (~2) which include water bodies, urban areas, and snow. While the MCD64A1 product provides a pixel-level burn-date uncertainty, to ensure fairness in the comparison we did not use this layer since the FireCCI51 product does not provide comparable information. For the spatial accuracy assessment, the pixels of both products were resampled (nearest neighbor) to a 20-m grid as dictated by the validation reference.
data (Section 2.2). Each 20-m resampled MCD64A1 and FireCCI51 burned grid cell was assumed to have 100% BA, and all resampled grid cells that overlapped with non-cropland areas of the reference imagery were excluded from our analysis.

2.2. Cropland burning reference areas

Seven reference areas within Ukraine’s croplands that cover almost 5% of Ukraine’s land area were manually digitized and classified to create highly detailed burned versus unburned cropland field maps.
fire product contains a number of variables including latitude, longitude, date, and UTC time. The daily AF count within each location’s mapping period was summed and compared to the daily mapped BA from the MCD64A1 and FireCCI51 products. The AF products offer both an accurate time of burning and are also able to identify smaller fires than the BA products (Giglio et al., 2003; Oliva and Schroeder, 2015).

2.4. Confusion matrix and accuracy metrics

We performed an assessment of the accuracy of the spatial patterns of BA by computing an error matrix to quantify the correspondence of the burned and unburned pixels reported in the two MODIS-based products with the gridded reference pixels. Each reference area grid cell has an associated BA percentage, therefore, a sensitivity analysis was performed to quantify the potential range of accuracy metrics within the analysis based on thresholding the reference BA percentage values for each reference area (A–G) and the combined seasonal outputs: Spring (A and B) and Summer (C–G). Several BA percentage threshold ranges were chosen for the sensitivity analysis: 1–100%, 25–100%, 50–100%, 75–100%, and 100% BA reference grid cells only. Since the accuracy metrics across the range of BA percentage thresholds did not vary substantially, we ultimately chose a reference BA percentage threshold of 50–100% to represent the minimum BA percentage per field because: i) the majority of reference field polygons within each reference region contain 50% or more BA and therefore give a representative sample, ii) a field (often only a few MODIS pixels in size) with less than half of its area burned (e.g., some fields contain only 1% BA as seen with 3-m Planet data) will more likely be missed by the MODIS-based BA products and therefore will put an unrealistic expectation on these coarse BA products, iii) the MODIS-based product pixels are assumed to contain 100% BA and therefore only including reference burn grid cells with 50% or more BA will help reduce the potential bias in commission error from the larger MODIS pixels, and iv) it reduces the impact of the inherent fuzziness in the smallest resolvable burn patch within our reference dataset given that the fields were classified using imagery having three different spatial resolutions (3 m, 20 m, and 30 m). Some fields, therefore, might be classified using the super-resolution of Planet (i.e. small burn patches only a few Planet pixels large) while others may be classified via the 20-m or 30-m moderate resolution imagery.

The error matrices were produced by comparing the area and proportions of agreement and disagreement between the burned and unburned pixels within the gridded 20-m reference pixels (minimum and maximum BA) and the resampled global BA products (Table 1). The “unburned” grid cells in the global BA products included values

Table 1
Confusion matrix elements describing the area (Aij [km2]) or the proportion (Pij) [0–1]

<table>
<thead>
<tr>
<th>Classified Data</th>
<th>Burned</th>
<th>Unburned</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burned</td>
<td>A11 (P11)</td>
<td>A21 (P21)</td>
</tr>
<tr>
<td>Unburned</td>
<td>A12 (P12)</td>
<td>A22 (P22)</td>
</tr>
</tbody>
</table>

Table 2
Unburnable (class −2) 20-m grid cells with an associated reference field grid cell. Units: km2 and percentage of total reference field grid cells.

<table>
<thead>
<tr>
<th>Region</th>
<th>FireCCI51 (km2)</th>
<th>MCD64A1 (km2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>8.5 (0.23%)</td>
<td>0.0 (0.0%)</td>
</tr>
<tr>
<td>B</td>
<td>8.5 (0.24%)</td>
<td>0.0 (0.0%)</td>
</tr>
<tr>
<td>C</td>
<td>3.3 (0.22%)</td>
<td>0.0 (0.0%)</td>
</tr>
<tr>
<td>D</td>
<td>4.9 (0.20%)</td>
<td>0.0 (0.0%)</td>
</tr>
<tr>
<td>E</td>
<td>8.1 (0.34%)</td>
<td>0.0 (0.0%)</td>
</tr>
<tr>
<td>F</td>
<td>3.6 (0.30%)</td>
<td>4.4 (0.36%)</td>
</tr>
<tr>
<td>G</td>
<td>7.1 (0.27%)</td>
<td>4.3 (0.17%)</td>
</tr>
</tbody>
</table>

2.3. MODIS and VIIRS active fire products

The 1-km MODIS (MCD14ML C6 V3; Giglio et al., 2016) and the VIIRS 375-m (VNPI4MGML CI V2; Schroeder et al., 2014) AF products were downloaded from the University of Maryland’s FTP server (ftp://fuoco.geog.umd.edu/Giglio et al., 2020). The AF points were buffered using a variable buffer radius (based on the location within the swath grid) to account for the variation in footprint across the MODIS and VIIRS scans. The AF points were visually analyzed and removed from the analysis if they were not associated with cropland burning. Each active

(42,958 fields in total) over a range of dates in 2016 and 2017 (Fig. 1). These reference areas were created via exhaustive visual interpretation of all available 20-m Sentinel-2 Multi-Spectral Instrument (MSI), 30-m Landsat-8 Operational Land Imager (OLI), and 3-m Planet imagery (www.planet.com), in conjunction with filtered Visible Infrared Imaging Radiometer Suite (VIIRS) AF point data (VNP14IMGML; Schroeder et al., 2014) and selected ground-based observations provided by a local agricultural expert. The mapping locations and dates were chosen to give a representative sample of the spring and summer crop residue burning seasons in Ukraine. Dates were chosen to represent either the first half of the burning season or the second half through plotting the distribution of VIIRS AF points throughout 2016 and 2017. For example, in region F, the peak of burning occurred on 27 July 2017, therefore, the mapping period was chosen from 1 June 2017 until the peak. Whereas, in region D, the peak occurred on 1 August 2016, therefore the mapping period was chosen to represent the burning that occurred from the peak until 31 August 2016.

Each digitized polygon was attributed with the following field classification: 1 = active flame/smoke or BA with corresponding VIIRS AF polygon (i.e. an overlapping polygon with a date aligned with the visual change on the field); 2 = definite BA but with no flame/smoke or AF point; 3 = ambiguous (a distinct darkening occurred on the field, but the analyst is unsure if the field was burned then plowed or only plowed); 4 = definitely unburned; 5 = not cropland or fields are too small that land cover conditions were difficult to determine on very high resolution (3 m) imagery (Fig. 2). Each Class 1, 2, and 3 field was also attributed with the fraction burned (see Hall et al., 2021 for details). For the sake of clarity, the classes will hereafter be referenced using the following naming convention: definite burn (Class 1 and Class 2), possible burn (Class 3), unburned (Class 4), and non-cropland/other (Class 5).

All regions had at least daily Planet imagery – in some regions both morning and afternoon overpasses were available – as well as the accompanying Sentinel-2 and Landsat-8 images. None of the regions experienced persistent cloud cover, therefore, analysts were able to view the fields almost daily during the mapping periods. Any instances of distinct darkening on the field where the cause could not be determined (i.e. timing of burning not coinciding with the overpass or due to cloud cover on the day of burning) resulted in a possible burn classification that we use to weigh the impact of our uncertainty. Although there is a larger proportion of reference fields within the possible burn class during the spring compared to the summer, the vast majority of reference fields (95% spring and 98% summer) do not fall within this ambiguous class.

Although 3-m Planet data was a primary source for the visual interpretation, the fields were hand-drawn using Sentinel-2 (20-m) imagery as the base map. Therefore, for this study, the digitized polygons were rasterized to a 20-m comparison grid for the validation assessment. Each 20-m grid cell was attributed with the appropriate class and the fraction burned (Class 4), and definite burn (Class 1 and Class 2), definite unburned (Class 3), ambiguous (a distinct darkening occurred on the field, but the analyst is unsure if the field was burned then plowed or only plowed); 4 = definitely unburned; 5 = not cropland or fields are too small that land cover conditions were difficult to determine on very high resolution (3 m) imagery (Fig. 2). Each Class 1, 2, and 3 field was also attributed with the fraction burned (see Hall et al., 2021 for details). For the sake of clarity, the classes will hereafter be referenced using the following naming convention: definite burn (Class 1 and Class 2), possible burn (Class 3), unburned (Class 4), and non-cropland/other (Class 5).

All regions had at least daily Planet imagery – in some regions both morning and afternoon overpasses were available – as well as the accompanying Sentinel-2 and Landsat-8 images. None of the regions experienced persistent cloud cover, therefore, analysts were able to view the fields almost daily during the mapping periods. Any instances of distinct darkening on the field where the cause could not be determined (i.e. timing of burning not coinciding with the overpass or due to cloud cover on the day of burning) resulted in a possible burn classification that we use to weigh the impact of our uncertainty. Although there is a larger proportion of reference fields within the possible burn class during the spring compared to the summer, the vast majority of reference fields (95% spring and 98% summer) do not fall within this ambiguous class.

Although 3-m Planet data was a primary source for the visual interpretation, the fields were hand-drawn using Sentinel-2 (20-m) imagery as the base map. Therefore, for this study, the digitized polygons were rasterized to a 20-m comparison grid for the validation assessment. Each 20-m grid cell was attributed with the appropriate class and the fraction burned (if applicable) using a maximum-area cell assignment rule. All grid cells assigned as non-cropland/other were removed from the analysis. Minimum and maximum reference BA, respectively represented by the definite burn class by itself, and the definite burn plus the possible burn classes in combination, were calculated by multiplying the total polygon/field area by the fraction burned. The subsequent analysis was performed within each region A–G individually and also grouped into seasonal outputs: Spring (A and B) and Summer (C–G).

2.3. MODIS and VIIRS active fire products

The 1-km MODIS (MCD14ML C6 V3; Giglio et al., 2016) and the VIIRS 375-m (VNPI4MGML CI V2; Schroeder et al., 2014) AF products were downloaded from the University of Maryland’s FTP server (ftp://fuoco.geog.umd.edu/Giglio et al., 2020). The AF points were buffered using a variable buffer radius (based on the location within the swath grid) to account for the variation in footprint across the MODIS and VIIRS scans. The AF points were visually analyzed and removed from the analysis if they were not associated with cropland burning. Each active
0 (unburned) and −2 (unburnable). The following error analysis follows the same conventions and notations as outlined within Boschetti et al., (2019). Since there were no unmapped (−1) regions within the validation areas, and the number of unburnable (−2) grid cells was extremely small (Table 2), the confusion matrices only consider the burned and unburned classes.

Error matrices were calculated for each region A–G separately and also combined seasonally. Five commonly used accuracy metrics (equations (1)–(5)) were calculated using the proportions calculated in Table 1 (Padilla et al., 2014, 2017; Boschetti et al., 2016, 2019): Overall Accuracy (OA), Omission Error Ratio (Oe), Commission Error Ratio (Ce), the Relative Bias (relB), and the User’s Accuracy for the burn category (UA_Burn).

\[
OA = P_{11} + P_{22}
\]  
\[
Oe = \frac{P_{21}}{P_{11} + P_{21}}
\]  
\[
Ce = \frac{P_{12}}{P_{11} + P_{12}}
\]  
\[
relB = \frac{P_{21} - P_{12}}{P_{11} + P_{21}}
\]  
\[
UA_{Burn} = \frac{P_{11}}{P_{11} + P_{21}}
\]

For illustration purposes, only the seasonal confusion matrices and accuracy metrics for reference grid cells with 50–100% BA values are shown in Section 3.3 (Tables 4 and 5).

3. Results

3.1. Unmapped & unburnable classes assessment

Both products supply a value to represent an unmapped class (−1) for the case of insufficient input data, and a separate value to represent water (−2). The FireCCI51 product broadens the water class (−2)
further to include land pixels deemed unburnable (e.g., urban areas).

A crucial first step in any assessment is to visualize the data, a process which facilitates the interpretation of results and helps identify data artifacts that may lead to incorrect conclusions. A visual examination of both products found large swaths of unmapped pixels (~1) within the FireCCI51 product that were often artificially cut off along MODIS tile boundaries (Fig. 3; top right) and meridians of longitude (Fig. 3; bottom right). Despite the large size of these artifacts, none of the seven Ukraine reference areas contained this unmapped class for either product. This same tiling artifact was recently highlighted by Liu and Crowley (2021) over several cropland regions, including Ukraine and southern Russia.

Analysis of the water/not burnable class (~2) found only a small percentage (Table 2) of the 20-m reference pixels allocated as a field were designated as unburnable by the two products. While the majority of these omissions occurred in regions F and G situated along the southern coastline, the FireCCI51 product mapped field pixels, which are in principal potentially burnable, as unburnable in every region. This characteristic is presumably the result of an overly broad definition of unburnable surfaces used within the FireCCI51 mapping algorithm.

<table>
<thead>
<tr>
<th>Region</th>
<th>Minimum Reference BA (Definite Burn + Possible Burn)</th>
<th>Maximum Reference BA</th>
<th>MCD64A1</th>
<th>FireCCI51</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>42(37)</td>
<td>78(66)</td>
<td>65</td>
<td>11</td>
</tr>
<tr>
<td>B</td>
<td>268(240)</td>
<td>364(324)</td>
<td>550</td>
<td>59</td>
</tr>
<tr>
<td>C</td>
<td>398(367)</td>
<td>424(392)</td>
<td>362</td>
<td>230</td>
</tr>
<tr>
<td>D</td>
<td>253(230)</td>
<td>313(287)</td>
<td>245</td>
<td>333</td>
</tr>
<tr>
<td>E</td>
<td>165(156)</td>
<td>202(189)</td>
<td>144</td>
<td>76</td>
</tr>
<tr>
<td>F</td>
<td>134(125)</td>
<td>154(143)</td>
<td>80</td>
<td>103</td>
</tr>
<tr>
<td>G</td>
<td>645(592)</td>
<td>704(649)</td>
<td>565</td>
<td>629</td>
</tr>
<tr>
<td>Total</td>
<td>1905(1747)</td>
<td>2239(2050)</td>
<td>2011</td>
<td>1441</td>
</tr>
</tbody>
</table>

Table 3
Total burned area (BA; km$^2$) per reference region. Total BA for reference fields with 50–100% BA are shown in parentheses.

Fig. 4. Location B (March 2017) maximum reference BA fields overlaid with MCD64A1 (left; yellow) and FireCCI51 (right; purple). For illustration, the reference area field polygons with 50% or more BA are represented in black and less than 50% in dark grey. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
Seasonal confusion matrices reported in area (km$^2$) for the two extremes of the reference imagery (Min. BA and Max. BA) as determined by the treatment of the possible burn reference class.

3.2. Total burned area in reference areas

We calculated and compared total BA for the two MODIS-based products and the minimum and maximum reference BA layers (all BA % and fields with >50% BA) for each reference region using the 20-m gridded layers. In total, the FireCCI51 product mapped 464 km$^2$ to 798 km$^2$ less BA than the minimum and maximum reference BA totals (306 km$^2$ – 609 km$^2$ compared to reference fields with >50% BA), respectively. Whereas, the MCD64A1 product mapped 106 km$^2$ more BA than the minimum reference BA and 228 km$^2$ less than the maximum reference BA (264 km$^2$ more BA than the minimum reference BA and 39 km$^2$ less compared to reference fields with >50% BA). However, when comparing the total BA per region, no distinct pattern seems to emerge between the two products (Table 3). For example, in regions A, B, C, and E, the MCD64A1 product maps more BA than the FireCCI51 product, whereas, in regions D, F, and G, the converse is true. We observed a particularly high omission of burned area for the FireCCI51 product in region B (Fig. 4), but for no reason that could be discerned from an examination of the product’s supplementary data layers.

Table 5
Seasonal accuracy metric ranges (expressed as percentage) derived using equations (1)–(5). The uncertainty range of the accuracy assessment is based on the inclusion of the possible burn class with the definite burn class (maximum) compared to only using the definite burn class (minimum).

Fig. 4. Location B (March 2017) maximum reference BA fields overlaid with MCD64A1 (left; yellow) and FireCCI51 (right; purple). For illustration, the reference area field polygons with 50% or more BA are represented in black and less than 50% in dark grey. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
3.3. Confusion matrix and accuracy metrics

Quantifying the total BA per region is useful for understanding the broader comparison between the two global products, but it does not assess the accuracy of the spatial patterns. Assessment of the correspondence between the burned and unburned pixels reported in the two MODIS-based products with the gridded reference pixels (BA > 50%) is illustrated through seasonal confusion matrices (Table 4) and seasonal accuracy metric ranges (Table 5).

Although the OA of both products is relatively high (>80%), this metric is heavily influenced by the product’s ability to map the unburned class. The Oe and Ce values are, in contrast, more informative of the products’ ability to map small burns. Here the Oe and Ce values were high for both products, a result that is not unexpected due to the nature of cropland burning. Equally unsurprisingly, the Spring Oe and Ce values were higher than the Summer values as the visual (3-m Planet imagery) and the spectral distinction between BA and plowed fields is harder to discern due to the wet (post-snow melt) field conditions (e.g., Hall et al., 2016). Specifically, in Ukraine, springtime cropland burning is associated with burning fields in preparation for planting corn/maize, whereas, in the summertime the burning is associated with post-harvest wheat residue burning (Hall et al., 2021; Korontzi et al., 2006). This visual and spectral change between a dry, fully-vegetated wheat field in the summer and a harvested (and then burned) field is more distinct and easier for the BA algorithms to detect, which is clearly represented by the overall lower summertime Oe and Ce values for both products (Table 5). Interestingly, although the MCD64A1 product has a fairly comparable Oe and Ce (more so for the Oe) between both seasons, the FireCCI51 product has a very distinctive increase in both the Oe and Ce values in the spring reference areas. The higher Oe in spring for the FireCCI51 product seems to be related to the unexplained lack of BA mapped within both spring reference areas at Locations A and B (Fig. 4).

The relative bias (relB) quantifies the degree to which the MODIS-based BA products underestimate (negative bias) or overestimate (positive bias) BA relative to the reference BA while allowing errors of commission to offset errors of omission. The relative bias of the two products was virtually identical for the summer burning season but showed substantial differences in the spring. During this season the largest bias for MCD64A1 occurred with respect to the minimum reference BA layer (definite burn class only), with MCD64A1 significantly underestimating the BA (~55%), while the largest bias for the FireCCI51 product occurred with respect to the maximum reference BA layer (definite burn plus possible burn) with a significant overestimation of 86%. While a lower product bias is almost always desirable, we caution that the relB error metric can be skewed through manipulation of the Oe and Ce scores. For example, adjusting a BA mapping algorithm to produce more commission errors (i.e., false alarms), even on a purely random basis, will reduce the magnitude of a negative relative bias. Conversely, removing burned pixels, even on a purely random basis, can lower a high relative bias. Finally, as with the Oe and Ce values, the user’s accuracy (UA_{burn}) — the probability that a mapped burn pixel represents a burn in the reference data — further highlights the inaccuracies of both BA products in spring.

3.4. Intercomparison between MCD64A1, FireCCI51, MCD14ML, and VNP14IMGML

We undertook a temporal intercomparison of the two MODIS-based BA products with the MODIS (MCD14ML) and VIIRS (VNP14IMGML) AF products to understand the temporal patterns and bias in relation to the known dates recorded in the AF observations. In particular we sought to determine if the FireCCI51 product also confuses the harvest signal which is a known problem within the MCD64A1 product (Giglio et al.,...
The analysis was split into three separate groups: spring early peak (regions A and B; Fig. 5), summer early peak (regions C, E, F, and G; Fig. 6), and summer late peak (region D; Fig. 7).

The harvest confusion in the MCD64A1 product (Fig. 6) is clearly seen at the start of the summer burn season with approximately two weeks of recorded BA (dark blue bars) before the start of the AF observations. Although there are a few instances of recorded BA early in the summer timeseries from the FireCCI51 product, the harvest signal does not seem to impact this product as it does the MCD64A1 product. A similar pattern was also seen in the spring (Fig. 5), however, the explanation for why the peak of BA occurs before the AF peak is not fully known. A potential explanation for this effect is related to confusion with snowmelt in these dark soil regions (Roy et al., 2005).

Fig. 6. Summer beginning (Region C, E, F, and G) daily MODIS-based burned area (km$^2$) and MODIS (MCD14ML) and VIIRS (VNP14IMGML) active fire counts over each mapping period (see Fig. 1).
Despite the densely-packed, large cropped fields and the full-field burns, both BA products suffered from high Oe and Ce values further supporting the findings found in previous cropland burning studies (e.g., Hall et al., 2016; Lasko et al., 2017). These BA mapping inaccuracies within our seven regions have a number of potential causes including harvest and snow-melt confusion. Some studies have also found performance differences between these two BA products within cropland areas to be related to the underlying land cover products used within each algorithm: MCD64A1 uses the MODIS MCD12Q1 land cover product (Friedl et al., 2010) while FireCCI51 uses the CCI-Land Cover Phase 2 product (Bontemps et al., 2015). Specifically, each algorithm applies different rule sets for burning within cropland, therefore these will be subject to any errors within the underlying land cover products (Giglio et al., 2018). For example, Campagnolo et al. (2019) found high MCD64A1 Oe values within the Russian cropland which were likely attributed to the fragmented cropland map within the MCD12Q1 product. Although in our seven mapped regions, we did not find any noticeable differences or cropland fragmentation within the underlying land cover products.

While the MCD64A1 and FireCCI51 burned area products have been extensively validated for wildfires and larger prescribed burns, our analysis is a first step toward introducing a comparable level of rigor for the more difficult case of cropland field burning. The intensely labor-intensive effort required to manually produce our BA reference maps necessarily limited this first assessment to CEOS validation Stage 1, whereby product accuracy has been “estimated using a small number of measurements obtained from selected locations and time periods” (http://lpvs.gsfc.nasa.gov/). As part of a recent CCI-sponsored effort, independent BA reference data produced under the CEOS protocol were compiled into the Burned Area Reference Database (Franquesa et al., 2020) to facilitate the development and validation of new BA products. Although the resulting compendium is extremely useful for validating BA maps of wildfires and many prescribed burns, cropland fire validation requires a different sampling strategy. Prior to this study, cropland-fire training and validation samples have almost universally been compiled through sporadic field-based samples (often taken along the road) or via surveys (e.g., Singh et al., 2021). Neither approach is entirely suitable for comprehensive accuracy assessments that include cropland. In Ukraine and Russia, for example, open-burning is illegal, and survey data are therefore often unreliable and biased toward farmers who do not burn their fields (Hall et al., 2016; SovEcon, 2013). Instead, comprehensive BA reference data produced specifically for cropland are needed to help algorithm developers improve their products within this land cover class. For example, these validation results (i.e. harvest and possible snowmelt confusion, tile-boundary seams, etc.) will be used to help guide the upcoming MCD64A1 Collection 7 burned area product improvements within cropland.

Over the past several years, new methods to improve the representation of small fires in global or regional emission inventories using MODIS or VIIRS AF data have been developed (e.g., Randerson et al., 2012; Singh et al., 2020; Shi et al., 2020; Yin et al., 2021). These methods attempt to exploit the fact that AF products can generally detect fires that are several orders of magnitude smaller than the sensor footprint (Giglio et al., 2003). Compared to other land cover types, cropland regions contain small, fragmented fires making AF data an attractive alternative for crop emission analysis (e.g., Hall et al., 2021; Liu et al., 2020b). The more accurate timing of AF observations can also provide additional constraints on commission errors, which for the MCD64A1 product can arise from confusion with the harvest signal. With the launch of new moderate-resolution (~20-m) sensors, there has been a move toward capturing small burns using a combination of Landsat and Sentinel imagery (e.g., Roteta et al., 2019; Roy et al., 2019). Despite the finer spatial resolution of these sensors, the ~3- to 5-day gap in overpass time (without taking into account cloud cover) will cause relatively high errors, especially in areas where farmers plow soon after burning, or if the spectral thresholds are not designed to avoid the harvest-signal confusion (Li & Roy, 2017; van Dijk et al., 2021).

5. Conclusion

We assessed the accuracy of two MODIS global BA data sets, NASA’s 500-m MCD64A1 burned area product and the ESA’s CCI 250-m FireCCI51 burned area product, in cropland (primarily wheat and maize) using a representative sample of agricultural sites in Ukraine during 2016 and 2017. BA reference maps were generated through exhaustive visual analysis of 3-m Planet, 20-m Sentinel-2, and 30-m Landsat-8 imagery in conjunction with 375-m VIIRS active fire observations in consultation with a local agricultural expert. Initial visual analysis of the MODIS-based BA products found large swaths of tiling artifacts in the unmapped class of the FireCCI51 product, however, this issue did not impact our seven reference areas.

We found that despite the large field sizes and more intensive cropland fires in Ukraine, both products suffer from very high errors of omission: MCD64A1 (71–76%) and FireCCI51 (63–99%). In all seven reference areas, the finer spatial resolution of the 250-m FireCCI51 product seemed to offer little or no advantage for mapping these small fires over the 500-m MCD64A1 product – a finding that has been found in several other studies (e.g., Fornacca et al., 2017; Vetrita et al., 2021). In addition, we undertook a temporal intercomparison of the MCD64A1 and FireCCI51 BA products with 1-km MODIS and 375-m VIIRS active fire products to understand the temporal patterns and bias in relation to the known dates recorded in the active fire observations. Here we will be used to help guide the upcoming MCD64A1 Collection 7 burned area product improvements within cropland.
confirmed the harvest confusion in the MCD64A1 product at the start of the
summer, predominately wheat, burning season.
Given the coarse-resolution (≥250 m) of the two global BA products,
Ukraine was an optimal location to conduct this cropland-focused BA
validation. The large field sizes and intense fires (primarily due to high
yields) compared to other crop-producing countries (e.g., India) pro-
vided an opportunity to validate and understand these products for what
is in many respects a “best-case” scenario. The ability to map cropland
BA using coarse-resolution sensors and methods based on spectral
reflectance will always be compromised due to the short-lived burn
signatures between the time of burning and the subsequent plowing
and/or seeding. The high omission and commission errors we found for
both products within a region having relatively advantageous charac-
teristics for cropland BA mapping (e.g., large fields) demonstrates that
considerable caution is required when using MODIS-based BA products
in cropland areas, especially those in which the fields are typically
smaller (e.g., South-East Asia). Since these BA products provide a pri-
mary input into fire emission databases and atmospheric transport
models, the large cropland BA omission errors will have detrimental
downstream impacts on the emission databases and subsequent sci-
entific and policy-related findings.

CRediT authorship contribution statement

Joanne Hall: Conceptualization, Data curation, Formal analysis, Methodology, Writing. Fernanda Arqueta: Visualization, Data curation, Writing - review & editing. Louis Giglio: Conceptualization, Funding acquisition, Methodology, Writing (original and review).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence
the work reported in this paper.

Acknowledgements

We thank Mr. O. Zhuravel (Food and Agriculture Organization of the United Nations, Ukraine and Ukrainian farmer) for his valuable local
knowledge of agricultural practices in Ukraine. This work was supported by

References

modelling: achievements of the Land Cover component of the ESA Climate Change
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


Singh, T., Biswal, A., Mor, S., Ravindra, K., Singh, V., Mor, S., 2020. A high-resolution emission inventory of air pollutants from primary crop residue burning over Northern India based on VIIRS thermal anomalies. Environ. Pollut. 266, 115132.


Zhang, T., Wooster, M.J., De Jong, M.C., Xu, W., 2018. How well does the ‘small fire boost’ methodology used within the GFED4.1s fire emissions database represent the timing, location and magnitude of agricultural burning? Remote Sens. 10 (6), 825.
