Assessing the Shape Accuracy of Coarse Resolution Burned Area Identifications

Michael L. Humber, Luigi Boschetti, and Louis Giglio

Abstract—Accuracy assessment of burned area maps has been traditionally performed using pixel-based metrics, with the objective of assessing the accuracy and precision of burned area estimates at local and regional scales. While these assessments are helpful for obtaining consistent estimates of the burned area across many fires and over large areas, pixel-based approaches do not necessarily characterize how well individual fires are mapped. At the individual fire scale, other factors like the size of the fire have significance regarding ecology, fire succession, and landscape management and determining other fire properties such as the spread rate. We propose a method for evaluating wildfire classification maps, which retains the spatially explicit properties of the burn scar. Our method quantifies the edge error (EE) of burned area classifications and reference maps by calculating the average geometric normal of the evaluated burned area boundary along the burn edge and the two nearest neighbor samples from the reference burn boundary. The metric is a physically meaningful quantification of the EE, which represents the average distance between the boundaries of the reference and evaluated burn scars. The methods are demonstrated by comparing MODIS Burned Area (MCD64A1) maps to Monitoring Trends in Burn Severity (MTBS) maps for 173 total wildfires in the United States. The results indicate that when accounting for the minimum achievable EE (MAEE) due to differing spatial resolutions, the mean EE is less than two MODIS pixels and the magnitude of the errors does not appear to be related to fire size.

Index Terms—Accuracy assessment, MODIS Burned Area, North America, wildfire.

I. INTRODUCTION

WILDFIRES are an important natural process which have widespread effects on human health, property, ecology, land cover, emissions, and more [1]. While the fire activity over the last millennium can be modeled effectively using climatological variables until the Industrial Revolution, increased pressure from the growing human population around the 1900s altered global fire patterns via anthropogenic suppression and ignitions [2]. This change is not limited to solely the number of fires and amount of burned area: recent studies have shown that fire size, and as a corollary, shape, are influenced by human activity as well [3], [4].

From a remote sensing perspective, understanding and mapping the burned area has received more attention in the past than understanding the size and shape of individual fires. In fact, current methods used for identifying individual fires from coarse resolution satellite data require extracting those fires from existing burned area maps [3]–[8]. However, the shape and size of individual fires is an important topic with regard to ecology and fire succession, landscape management, and determining other fire properties such as the spread rate. As a basic example, the length of the fire front impacts the ability of fauna to escape the flames [9]. The postfire succession can be influenced by the shape of the burn as well as by the patchiness of the burned area mosaic which can favor certain plant or animal traits by changing the amount of fringe habitat and the openness of the canopy [9], [10]. Fire size is also related to management practices—heterogeneous landscapes create fuel breaks which can limit the spread of fire across the surface [11], [12]. Finally, in image processing workflows such as those presented in [8], [13], and [14], individual fires are identified for the purpose of extracting other metrics such as the spread rate, which are inherently linked to the shape and size of the fire.

Several programs exist with the goal of providing satellite products to be used for operationally monitoring—spatially and temporally—global wildfire activity. Such products can be broadly categorized as active fire products (representing locations actively burning on the Earth’s surface at the time of the sensor overpass) and burned area products (representing the postfire-affected area as determined by the removal of vegetation, exposure of soil, and the presence of charcoal and ashes). Two decades of mapping efforts have produced a number of global coarse spatial resolution (e.g., 250-m to 1-km pixel size) burned area products, including MCD45A1, MCD64A1, Copernicus Burnt Area, Fire CCI, and others (respectively, [15]–[19]). These products have used input from a variety of sensors including MODIS, SPOT-VEGETATION, PROBA-V, and MERIS. The extent and timing of burning is an essential parameter in fire emission calculations performed with the conventional bottom-up approach [20], and the need for consistent estimates of greenhouse gas emissions was one of the main drivers of the development of global satellite fire monitoring products [21].

Quality assessment of coarse resolution burned area products is needed to provide data users with necessary information about the suitability of the products for specific applications and has taken many forms such as intercomparison with other coarse resolution burned area or active fire products [22]–[26].
or comparison with a sample of higher resolution, independently derived reference burned area maps [27]–[30]. Product validation is an important activity outlined by the Committee on Earth Observation Satellites (CEOS) Land Product Validation (LPV) Subgroup and involves assessing product accuracy in one of four stages, each with increasing statistical rigor. The comparison with independent reference burned area maps (commonly termed validation) is conventionally conducted using accuracy metrics derived from a confusion matrix—i.e., the matrix reporting the co-occurrence of proportion of burned and unburned data in the product and in the independent reference data (for a review, see [31]), or from the regression between proportions of area burned in coarser resolution grid cells [32].

Arguably because of the great emphasis placed on the use of global burned area products for emissions estimation, validation has traditionally described the accuracy and precision of areal estimates at different scales, rather than the accuracy of other aspects of the burned area representation. The accuracy of the shapes mapped in a burned area product is currently not considered as part of product validation exercises, and neither is the accuracy of derived metrics such as fire size distribution, compactness, spread rate, and orientation of the fire front. There is an outstanding need to expand the validation of burned area products to consider these characteristics. Several recent studies have aimed to quantify the distribution of fire sizes and other characteristics of individual fires, based on existing data sets that have not been validated beyond standard areal accuracy [3], [4], [8], [33]. Of those studies, only the results of [3] were partially validated, limited, however, to the comparison of the size distribution of MODIS-derived burn scars to the size distribution of a sample of Landsat-derived burn scars, without directly comparing individual fires.

Evaluating properties such as the number of fires, fire size, and fire shape require object-based approaches, rather than area-based approaches. While object-based accuracy assessments have been previously applied to remotely sensed thematic maps (see [34]–[37]), there has been relatively little research on the applications in burned area detection. Early work on object-based accuracy assessment of burned area classifications was conducted by Remmel and Perera [38], who considered the degree of areal overlap between mapped and reference data between individual fire events using AVHRR-derived burned area maps and a wide variety of high-resolution reference data (in addition to the confusion matrix). While this work was based on the overlapping area and did not explicitly take into account the fire boundaries, it does highlight the errors from the perspective of the mapped burn, the individual fire event, and the reference data which provides an analogy to the concept of “producer’s” and “user’s” accuracy. Another recent exception is [13], who proposed a patch-based burned area product accuracy assessment approach, but their method approximates it with an ellipsoidal model for the purposes of compatibility with the behavior of more advanced vegetation and fire models, such as “Organising Carbon and Hydrology In Dynamic Ecosystems” (ORCHIDEE) rather than assessing the actual mapped shape of the fire complex [39], [40].

In this article, we provide a novel edge error (EE) metric which is used to quantify the degree to which coarse resolution burned area maps retain the shape of burn scars identified at moderate resolution, in keeping with established protocols for burned area product validation. The metric is demonstrated through a comparison of the MODIS MCD64A1 burned area product, which has a nominal resolution of 500 m [16], to the Landsat-based Monitoring Trends in Burn Severity (MTBS) products, which have a resolution of 30 m [41]. A calculation of the minimum achievable EE (MAEE) metric, which accounts for differences in pixel size, is detailed in Section III along with other object-based metrics from the literature. Section IV presents the performance of the metrics, and this article concludes with a discussion of the implications of implementing coarse resolution burned area products for representing individual fire shapes.

II. DATA

A. MCD64A1 Burned Area Product

Coarse resolution sensors such as MODIS provide global coverage with short revisit times (e.g., daily). This is advantageous for burned area mapping as the high temporal frequency improves the probability of obtaining cloud-free observations and can be exploited to more accurately determine the day of burning. On the other hand, such sensors are unable to capture fine details in the shape of objects on the ground due to their low spatial resolution, and in the case of burned area mapping, the minimum fire size which can be reliably mapped is larger than that obtainable by moderate resolution counterparts [42]. In this article, the latest Collection 6 MODIS MCD64A1 burned area product [16], [30] was selected because it is operational, global, and publicly available. The Collection 6 MCD64A1 product detects the most total burned area of any current operational product at coarse spatial resolution [26], including significantly more burned area than the previous Collection 5.1 MCD45A1 product [15], [23], with yearly global burned area increasing by approximately 26% [16].

The MCD64A1 burned area mapping algorithm combines daily MODIS surface reflectance imagery with 1-km MODIS active fire data to map burning on a daily basis at the MODIS 500-m spatial resolution. The algorithm applies dynamic thresholds to composite MODIS Terra and Aqua imagery generated from a burn-sensitive spectral band index derived from MODIS 1240- and 2130-nm Terra and Aqua bands, and a measure of temporal variability. Cumulative MODIS 1-km active fire detections are used to guide the selection of burned and unburned training samples and to guide the specification of prior burned and unburned probabilities [16].

The MCD64A1 burned area product includes several data layers—“Burn Date,” “Burn Date Uncertainty,” “QA” (Quality Assurance), and “First Day” / “Last Day” (during which burns can be reliably detected) [43]. The product is distributed in the MODIS Sinusoidal Equal Area Projection [44], with a nominal 500-m resolution (the actual resolution is 463.3127 m).
TABLE I
EIGHT CASE STUDY FIRES SELECTED IN THIS STUDY

| Fire Name                | State       | Year | Ignition Date | Size (ha) | Dominant Land Covers                                  |
|--------------------------|-------------|------|---------------|-----------|------------------------------------------------------|
| South Sarpy Fire         | Montana     | 2015 | 12-Jul-15     | 2,017     | Shrub/Scrub and Grassland/Herbaceous                 |
| Lincoln Canyon Complex   | New Mexico  | 2011 | 2-Aug-11      | 4,048     | Shrub/Scrub and Grassland/Herbaceous                 |
| Dry Creek Complex        | Washington  | 2009 | 20-Aug-09     | 20,170    | Shrub/Scrub                                          |
| Esmerelda Fire           | Nevada      | 2005 | 15-Jul-05     | 40,920    | Shrub/Scrub                                          |
| Rim Fire                 | California  | 2013 | 17-Aug-13     | 104,040   | Evergreen Forest and Shrub/Scrub                    |
| Cave Creek Complex       | Arizona     | 2005 | 21-Jun-05     | 108,605   | Shrub/Scrub                                          |
| Murphy Complex           | Idaho       | 2007 | 21-Jul-07     | 229,628   | Shrub/Scrub                                          |
| East Amarillo Complex    | Texas       | 2006 | 12-Mar-06     | 240,139   | Grassland/Herbaceous and Shrub/Scrub                |

Fires selected for this study, ordered by size (area reported by MTBS). Dominant Land Cover determined by the degree of overlap with the National Land Cover Database 2011 (NLCD).

B. MTBS

In the previous literature, coarse resolution (≥250 m) satellite-derived burned area maps have been assessed or validated using moderate (typically ≤30 m) resolution data such as those provided by Landsat (see [16], [27], [30], [31], [45]). There are multiple programs which map burned area across the conterminous United States and Alaska using Landsat data. One of the more comprehensive efforts with regard to the total number of fires mapped is the MTBS [41] project, which provides wall-to-wall Landsat-based burned area maps for the United States. The classification is largely derived from photointerpretation conducted by expert interpreters rather than automated methods.

MTBS commenced in 2005 in support of the Wildland Fire Leadership Council (mtbs.gov). Now supported by the USGS, U.S. Forest Service, and the U.S. Department of the Interior, the program aims to map all fires since 1984 which exceed 1000 (405 ha) or 500 acres (202 ha) in the western and eastern United States, respectively. Three basic types of data are available from the MTBS program: burned area boundaries, fire occurrences, and burn severity mosaics. Both the burned area boundaries and burn severity mosaics provide information about the location and spatial extent of fires occurring in the United States and selected territories.

The burned area boundaries data set consists of vectors which delineate the outermost extent of the burned area patches. The boundaries are derived via photointerpretation of Landsat TM, ETM+, and OLI scenes and do not identify internal unburned islands within the boundary of the burn [41]. The burn boundaries are used to limit the extent of analysis for the burn severity data, which consists of classifications derived from the pixel values indicating the severity of burning based on the differenced normalized burn ratio (dNBR).

For studies in the United States, MTBS data have been used as a reference data set for comparison to other products [46], [47]. However, studies have demonstrated that MTBS often overestimates the total burned area due to the commission of the unburned islands to the burned area total [48]. This feature is consistent with the intended use of MTBS, which focuses on land management rather than burned area estimates [41]. It is noted that due to the ambiguity in the burn severity classification, it is impossible to reconstruct internal unburned islands in the context of this study.

C. Study Area

Eight fires in the western United States, occurring between 2005 and 2015, were selected as case studies (Fig. 1 and Table I). The fires were selected in a semirandom fashion from the MTBS data set, such that the fires represented a variety of sizes and locations. No more than one fire was selected for any given state. According to the National Land Cover Database (NLCD2011) [49], the dominant land cover for the Dry Creek Complex, Esmerelda Fire, Cave Creek Complex, and Murphy Complex was shrub/scrubland. The South Sarpy Fire and Lincoln Canyon Complex were also predominantly in shrub/scrublands, but also included grasslands/herbaceous areas. The East Amarillo Complex, the largest fire in the study, occurred predominantly in grasslands/herbaceous areas with
Secondary occurrence in shrub/scrublands. Finally, the Rim Fire occurred predominantly in evergreen forest with a secondary land cover of shrub/scrublands. In addition to these eight fires, 165 fires identified from the 2016 burning season were selected to demonstrate the methods over a large sample size.

III. METHODS

Previous work has shown that the accuracy of pixel labels at the regional and continental scale does not necessarily indicate accuracy with respect to assigning shape boundaries [36]. Object-based approaches to assessing burned area detection accuracy, wherein the entire shape of a fire are taken into consideration rather than simply the individual pixels, are necessary to quantify a classifier's performance at the individual fire scale and enable the accuracy of the fire to be described with regard to shape as well as area. This approach, which should be considered complementary to—that rather than a replacement for—the commonly implemented pixel-based approaches, consists of three key steps: extraction and harmonization, metric calculation, and identification of the MAEE. These steps are described hereafter.

A. Extraction and Harmonization

Individual fires were extracted from their respective data sets using a two-pass region (otherwise known as “connected components”) labeling algorithm, such as that described in [50]. First, a binary mask was created from the MTBS and MCD64A1 data sets. The MCD64A1 product is distributed as a monthly composite; for this study, in instances where the burning event took place over the course of multiple calendar months, the monthly products were composited temporally such that the maximum day of burning between two consecutive months was retained. Unlike MTBS, the MCD64A1 product does not associate individual fires with a fire name. While several algorithms exist for the purpose of extracting individual fires from the MODIS Burned Area data sets (see [5], [7]), the operation was trivial and conducted manually for the relatively simple cases in this study. A binary mask was then created encompassing all cells flagged as burned.

From the binary masks, the locations of edge (that is, the boundary or perimeter) cells were extracted by identifying any cell adjacent to an unburned cell, based on queen’s case adjacency (otherwise known as 8-adjacency) rules. The location of the center of the cell was recorded, rather than the cell corners, and stored in a vector format. The boundaries of the MTBS fires were projected from the native Albers equal area projection to the MODIS Sinusoidal projection. For fires observed in 2016, MCD64A1-derived fires were then paired with MTBS fires under the following conditions: the overlapping area was greater than 10% of the MTBS and MODIS fire area; the area of the fire was greater than 500 ha; the fire was characterized as a “wildfire” by MTBS; and there was no obvious mischaracterization resulting from the rudimentary extraction method based on visual inspection.

B. Metric Calculation

1) EE Computation: Computer vision algorithms identify the similarity of two image objects through the lens of “shape representation” or “shape matching.” Generally, such algorithms may be used for database retrieval or image object retrieval [51]–[54]. Shape matching algorithms are not typically spatially explicit and instead focus on identifying patterns regardless of size or orientation [55], [56]. These features may be useful for identifying broad patterns of shape, but, for object comparison in the spatially explicit geographic domain, these may not be desirable attributes as the rotation or orientation of a fire scar on the landscape is an intrinsic property of the fire itself. Any agreement in burned area shapes along different orientations is, in this regard, coincidental.

An advantage of shape matching in the scope of this study is the ability of the algorithms to assess the similarity of object boundaries without the use of a user-defined parameter. The discrepancy in boundary locations, or so-called “contour dissimilarity,” is calculated by identifying the edges of an object (i.e., the burn edge extraction step) then calculating the distance between the edge locations of the evaluated object and a reference object. The error for each edge location is calculated by advancing through the edge points in order to identify the minimum distance between the objects [55].

Measures of contour dissimilarity are desirable in this regard because, assuming the data are represented in a projected coordinate system, the unit of the contour dissimilarity in the geospatial domain is a physically meaningful representation of distance. In this work, the contour dissimilarity is calculated based on EE, where the average EE represents the expected distance between a given evaluated and target object. The proposed EE metric quantifies the degree to which two burn identifications agree upon the location of a burn boundary. The method is used to determine the location of every edge pixel in an evaluated burn, Burn(eval), relative to its nearest neighbors (NNs) belonging to the target burn, Burn(tgt) (see Fig. 2). It is assumed that if the boundaries are closer together on average, then the representation of the burn shape as a whole is more accurate.

In the ideal case, zero EE represents instances where the burn boundaries of the evaluated product are perfectly aligned with the burn boundaries of the target product. In practice, this is very unlikely to be the case for an entire burn, especially at differing spatial resolutions, due to imperfect co-registration, subpixel differences in boundary identifications, and differences in methodologies for identifying burns in each data set. Note that while the first two issues are related to cell size and are not truly errors, the latter is a result of erroneous classifications. It follows that smaller EEs (those approaching zero), therefore, represent a higher level of agreement and a more accurate classification of the fire boundary while increasing EEs indicate poorer characterization of the fire boundary.

The (coarse resolution) MCD64A1 burn boundary is designated as the burn to be evaluated, Burn(eval), which is compared to the higher resolution MTBS burn boundary designated as the target burn, Burn(tgt). The EE is the mean error between an evaluated edge location Burn(eval) to the minimum of geometric normal of the line segment (⊥)
connecting the two NNs in a target burn, Burn(tgt)\(_{NN1}\) and Burn(tgt)\(_{NN2}\), or the closest of the two NNs such that for a Burn(eval) with \(n\) edge locations (Fig. 2)

\[
EE = \sum_{i=1}^{n} \frac{EE_i}{n}
\]

(1)

where

\[
EE_i = \min(EE_{NN}, EE_{Norm})
\]

(2)

and

\[
EE_{NN} = d(Burn(eval)_{i}, Burn(tgt)_{NN1})
\]

(3)

and

\[
EE_{Norm} = d(Burn(eval)_{i} \perp [Burn(tgt)_{NN1}, Burn(tgt)_{NN2}] )
\]

(4)

As detailed above, the value of EE (1) is the average of all EE\(_i\) [see (2)–(4)] for a given burn identification. That is, all edge cell locations of the evaluated burn are iteratively compared to the nearest edge cell location(s) of the target burn. In the event that a Burn(eval) has multiple Burn(tgt)\(_{NN}\) (when there is a tie for the second NN), EE\(_{Norm}\) is evaluated for all possible combinations of the tied elements, selecting the minimum of the evaluated outcomes. An example illustrating the EE\(_i\) vectors for a hypothetical pair of burn shapes is provided in Fig. 3.

It is noted that while a subset of the edge locations is often sampled in shape-matching implementations, all edge locations are selected in this methodology in order to retain the spatial integrity of the input data. To accommodate the analysis of this volume of data, the search for NNs is made more efficient (with respect to time) through the use of K-dimensional trees (“K-D trees”) \[57\]. K-D trees are a form of binary tree that can be used to rapidly reduce distance-based query time by dividing space using a hyperplane at each tree node. The time complexity for searching a K-D tree is approximated by dividing space using a hyperplane at each tree node. The operations needed to identify a desired value and the number of elements to be evaluated. For both the evaluated and reference objects, a K-D tree is constructed containing all points identified in the edge location extraction step. Each point along the edge of the test object is used to query the reference K-D tree to find the two NN points determined by the minimum Euclidean distance between the edge point and the point under evaluation in the K-D tree.

As the EE metric consists of the average distance between analogous points along the contours of two burned area identifications, the metric has physical significance and does not rely on any free parameter as input.

2) Computation of Overlapping Area Metrics: Many object-based metrics have been proposed in the literature, which relate the accuracy of a given evaluated object to a reference object based on area. Two of the more widely implemented indices, oversegmentation (OS) and undersegmentation (US), can be considered analogous to errors of omission and errors of commission, respectively \[34\]–\[37\]. OS describes the degree to which an algorithm divides an object into too many segments, i.e., commits areas which are outside the boundaries of the true object [34]. Thus, OS (5) defines the relationship between the overlapping area, or intersection, of the target object (“\(x\)”) and the evaluated object (“\(y\)”) to the area of the target object such that

\[
OS = 1 - \frac{\text{area}(x \cap y)}{\text{area}(x)}.
\]

(5)

Similarly, US (6) defines the relationship between the overlapping area of the target object (“\(x\)”) and the evaluated object (“\(y\)”) to the area of the evaluated object such that

\[
US = 1 - \frac{\text{area}(x \cap y)}{\text{area}(y)}.
\]

(6)
Both metrics were calculated assuming the MTBS burned area as the target object and the MCD64A1 burned area as the evaluated object. It is noted that while MTBS is designated as the target object (by convention) in this case, the MTBS burned cell identifications themselves are unvalidated and are expected to overestimate the total area burned due to the ambiguity of the “Unburned to Low Burn Severity” class. Hereafter, OS and US are also referred to as “overlapping area metrics” as they relate the area of one object to another.

3) Identification of the MAEE: While the EE is a measurement of the physical distance between comparable edge locations between two burned areas, the metric can only be interpreted directly when the burned areas are spatially co-registered and at the same spatial resolution. When the observations are presented at different spatial resolutions, it is necessary to account for the effects of the difference in spatial resolution in order to minimize the effects of random placement of burned cells in the higher resolution map compared to the lower resolution map.

In practice, many arrangements of burned pixels at high resolution can be accurately represented by a coarse resolution map such that the measured error is a consequence of the discrepancy in cell sizes. Fig. 4 illustrates two cases where a coarser resolution pixel accurately and reasonably preserves the shape of an object also represented at a higher resolution. In these cases, errors in the calculation of the EE are, therefore, the result of subpixel variations in the shape which are not expected to be captured by the coarse resolution product, rather than errors in the classification itself. Noteworthy, however, is that Fig. 4(a) demonstrates an EE close 0 according to the method described earlier (because all edge cells at the higher resolution intersect the NN or are located on the line segment connecting the two NNs, even though the areas are different!), while Fig. 4(b) exhibits a larger EE, though the value is less than the one-sided dimension of the coarse cell.

Calculation of the MAEE gives context to the measurement by providing an estimate of the unavoidable error which results from differences in spatial resolution rather than algorithm misclassification. The MAEE calculation is a simplified version of the method implemented by Boschetti et al. [58] for calculating the Pareto boundary. As with the cited work, it is necessary to have only the higher resolution image—the MTBS burned area map—and to know the cell size of the coarser resolution product (463.3127 m in the case of MCD64A1).

For each of the eight fires in this study, the MTBS maps were projected and resampled to the MODIS Sinusoidal 500-m grid such that the values of the output raster are a soft classification representing the proportion of the coarse cell that was identified as burned in the original map (values range from 0% to 100%; Fig. 5). Cells with values of 100% represent the core burned area, while locations near the perimeter of the fire exhibit decreasing burn proportions. The soft classifications were then hardened for all whole percent thresholds in the range [1%, 100%], resulting in 100 possible classifications for each fire. The extraction procedure was repeated on the hardened burn proportion maps (Fig. 6), upon which EE was calculated using the native resolution MTBS maps and the thresholded (MODIS resolution) MTBS proportion maps. Note that in cases where the cell sizes are the same and the data sets are spatially co-registered properly, the soft classification will contain only two unique values, 0% and 100%, where a threshold of 0% results in an (implausible) map where all cells are burned and a threshold of 100% results in the original burned area map itself.

The minimum of each EE series per fire represents the optimal, or most efficient, solution and is retained as the MAEE. Recalling that this number represents the amount of expected or unavoidable error due to the random placement of burned cells at the higher resolution relative to the coarse resolution, MAEE is reported along with EE in order to help distinguish between the errors resulting from incorrect classification from the errors resulting from differences in spatial resolution.

MAEE was calculated for only the eight case study fires. The reason for this is twofold: MAEE is unlikely to be relevant to users of the product when presented in aggregate, and as a practical matter, the calculation of MAEE is computationally intensive.

IV. RESULTS

The results of the methodology are presented in Sections IV-A and IV-B. The MAEE calculation procedures are presented first, as these results are a component of the final EE statistic. Then, the EE metric results are presented, followed by the overlapping area metrics. EE and overlapping area metrics are presented in aggregate for the 2016 fire season.

A. MAEE

The MAEE calculation was performed for each of the eight case study fires, taking into consideration the error from the MTBS at the MODIS 500-m resolution to the MTBS native resolution edge. For each fire and threshold for subpixel fraction of area burned (1%–100%), the mean EE is plotted in Fig. 7. Generally, the mean EE distribution is concave, which is to say EE decreases monotonically as the threshold increases until the minimum is reached, at which point the mean EE increases monotonically.

The minimum mean achievable EEs were observed using a minimum threshold in the range 11% (Murphy Complex
and East Amarillo Complex) to 29% (South Sarpy Fire). The MAEE ranged between \(\sim 120.68\) (South Sarpy Fire) and \(\sim 153.35\) m (Murphy Complex); thus, in all cases, the minimum mean achievable EE is less than 155 m, or roughly 33\% of a MODIS 500-m cell. The range of MAEE values and thresholds for each fire underscores the need to calculate the metric on a per-fire basis, rather than assuming a single global value.

### B. EE and Overlapping Area Metrics

The edge error metric, EE, is presented for each of the eight case study fires, where MTBS at the native resolution was used in all cases (not to be confused with the aggregated classifications used for calculation of the MAEE). The results are presented in Fig. 8 and Table II, which show that the EE is less than or equal to the MODIS cell size (461.3127 m) in 5 out of 8 possible cases and is slightly greater than the MODIS cell size in one other case—the South Sarpy Fire (466.34 m). The Rim Fire produces arguably the worst result, with EE exceeding 776 m or roughly 1.7 cell widths. For the 2016 fire season, the 25th, 50th, and 75th EE quantiles were 259.0, 332.9, and 442.7 m, respectively.

The EE does not appear to be driven by fire size for the fires in this study, indicating that the MCD64A1 detections along the edge of a fire are relatively stable. The highest edge accuracy was achieved by the Dry Creek Complex, the third smallest fire in the study which burned 20 170 ha according
to MTBS. The lowest edge accuracy was achieved by the Rim Fire, the fifth-largest fire in the study at 104 040 ha burned.

With regard to the Rim Fire, it appears that the burned area is poorly characterized by both the MCD64A1 and MTBS products. The former appears to omit some areas that were burned, while the latter commits a significant amount of burned area due to the ambiguous “Unburned to Low” burn severity class. This is most evident in the northern portion of the fire, as depicted in Fig. 9.

A comprehensive listing of the EE metric for the case study fires is provided in Table II, along with the overlapping area metrics. For both the OS and US metrics, the smallest fire in the data set, South Sarpy Fire, demonstrated the worst performance. Intuitively, the area-based metrics are highly susceptible to large swings in the value of the metric for smaller sample sizes. On the other hand, two of the larger fires, the Cave Creek Complex and Rim Fire, demonstrated the smallest OS and US, respectively. In the case of the Rim Fire, the US metric performed well at the expense of the OS metric for the reasons described above and illustrated in Fig. 9. We note that these results are in line with previous findings by Rodrigues et al. [59], who determined that the levels of OS and US (referred to as OEEDGE and CEEDGE, respectively) decrease as the average fire size increases for the Brazilian Cerrado. Regarding the 2016 fire season, the 25th, 50th, and 75th quantile OS errors were 0.15, 0.25, and 0.39, respectively, and the US errors were 0.04, 0.07, and 0.13, respectively.

V. DISCUSSION AND CONCLUSION

Burned area accuracy assessment has historically been limited to the traditional pixel-based confusion matrix approaches, which succinctly summarize the probability of a pixel having a correct burned or unburned label. These approaches are effective for studies related to the total area burned at coarse spatial scales where the actual shape of fires may be either obscured due to pixel size or of little consequence to the intended use of the data.

This article introduces a method for characterizing the accuracy of the shape of coarse resolution burned area detections, by comparing them to higher resolution reference burned area maps. A novel edge error metric (EE) indicates the average distance between the boundary of individual burned areas as mapped in the coarse resolution product and the reference boundary. This metric is accompanied by an indication of the MAEE which accounts for burning in the high-resolution reference map which is smaller than the resolution of the coarse product.

To benchmark the performance of the proposed metric, two conventional indices from the object-based image analysis literature were calculated—OS and US. The OS and US metrics show a general tendency to demonstrate large errors for small fires (consistent with [59]), while the EE does not appear to be related to fire size. This is intuitive given the formulation of the metrics—OS and US are indices based on the errors in 2-D area, while EE is a measure of the errors in zero-dimensional (point) edge locations. At the individual fire level, the proposed EE metric complements, but should not replace, the area-based indices because the quantities which they evaluate are different.

The approach was demonstrated by assessing the shape accuracy of the MODIS Collection 6 MCD64A1 burned area product, using as reference data a sample of high-resolution fire perimeters provided by the MTBS project. Our results indicate that for the sample of eight individual fires considered in the analysis, the MODIS Collection 6 MCD64A1 burned area product is able to capture the boundaries of fires identified at the Landsat-scale by MTBS. In most cases, the EE was less than the width of one MODIS 500-m cell and in all cases, the error was less than the width of two cells. Considering the sample of 165 fires that occurred in 2016, the average EE for the selected fires was approximately 332 m and the EE was less than two pixels in 160 out of 165 cases (97%). No anomalous/unanticipated algorithm behavior was observed when calculating the EE of the larger data set, indicating that our proposed method is operationally stable (a primary rationale for assessing the larger test sample).

It is important to note that the primary purpose of the MTBS data set is to provide information for land managers on burn severity, and it is not designed to be a reference data set for satellite-based burned area mapping [41], as the data set is often employed. As a result, and due to the labor-intensive procedure used to generate the MTBS data set, the extent of unburned islands within a fire is not mapped. However, it is expected that the identification of unburned islands within fires changes significantly with scale, e.g., many small unburned islands may exist at the 30-m scale which do not manifest as a meaningful signal at coarse resolutions. This highlights implicit assumptions of the EE metric that: 1) any boundary—burned or unburned—is large enough to be captured by both the test and reference data sets and 2) for any boundary in a given data set, a corresponding boundary exists in the other. In the absence of these conditions, the EEs increase as a function of the number, size, and placement of the unmatched boundaries within the fire extent. The unburned islands are discussed in the Appendix, with the EE metric applied to photointerpreted data to demonstrate the metric conceptually, assuming the aforementioned issues have been resolved.
Due to the cross-resolution calculation of the MAEE metric, this approach can theoretically be applied to any other combination of burned area maps and reference data, including maps and reference data of the same spatial resolution. In addition, given a method for automated extraction of individual fires from burned area maps (see [8], [60]), the method can conceivably be routinely applied to the global validation of coarse resolution products (see [28], [30]). The EE metric is able to capture information lost in pixel-based accuracy assessment and, for studies focused on shape and topology of burned area perimeters, it is a more representative and relevant source of information regarding mapping accuracy.

In future work, we will apply the EE metric as well as the other object-based indices in two research areas. First, we plan to analyze the difference in EEs for fires identified by different classification algorithms to compare, for example, the MODIS MCD64A1 Collection 6 500-m burned area product [16] and the Fire CCI version 5.1 250-m burned area product which is also derived from MODIS [61], or the Fire CCI version 4.1 300-m burned area product derived from MERIS [18].

In addition, given a higher spatial resolution data set delineating true fire boundaries, the EE metric can be used to refine burn scar extraction algorithms. Typically, individual burn scars are identified using flood fill algorithms which use a threshold for the maximum number of days between detections in neighboring pixels to determine adjacency (see [3]–[6], [8], [60], [62]), which can lead to US or OS if the threshold is too large or small, respectively. EE, in these cases, could be evaluated iteratively with different thresholds to empirically derive a more representative regional threshold.

**APPENDIX**

An implicit assumption in calculating the EE metric is that each object identified at the coarser resolution has a corresponding object at the finer resolution (this assumption is already made when selecting the fires for analysis in the first place). The EE can, therefore, be calculated including unburned islands within the outer boundary of the burn with no modification, provided this condition is met.

However, special consideration must be given to unburned islands as several studies have demonstrated that the presence of unburned islands observed by satellites do not correspond well to in situ measurements taken at a higher spatial resolution [63], [64]. Similarly, unburned islands identified at a higher resolution will not necessarily correspond to those identified at a coarser resolution, e.g., Landsat versus MODIS, meaning that it is not guaranteed that all unburned islands will be present in both data sets. This issue can, in theory, be solved by applying a robust set of rules for selecting unburned islands that are present across different resolutions, i.e., accounting for the low-resolution bias [58]. These rules could be based on the size of the unburned island, whether or not the unburned islands overlap in both data sets, etc. Determining these specific rules is outside of the scope of this work, but it is of interest to the broader fire remote sensing community and deserves additional study.
Two of the case studies were chosen to conceptually demonstrate the EE metric when applied to burn scars containing unburned islands. These fires, the Esmerelda Fire and the East Amarillo Complex, were manually photointerpreted to append the larger internal unburned islands to the MTBS boundaries (Fig. 10). The edge of the internal unburned islands and external fire boundaries were then compared between the modified MTBS and MCD64A1 fire shapes. The EE for the Esmerelda Fire was 252.59 m and the East Amarillo Complex was 398.52 m, representing improvements of more than 100 and 40 m, respectively, when compared to the EE of the fire boundary alone. This improvement can be attributed to the fact that the geometry of the unburned islands for the two burns is less complex (often elliptical in shape) than the fire boundary geometry. A caveat to this demonstration is that the photointerpretation of the unburned islands purposefully left out very small unburned islands—as previously mentioned, conducting this analysis at scale requires a robust method for determining the minimum detectable unburned island size.

References

[1] D. M. J. S. Bowman et al., “Fire in the Earth system,” Science, vol. 324, no. 5926, pp. 481–484, Apr. 2009.
[2] O. Pechnoy and D. T. Shindell, “Driving forces of global wildfires over the past millennium and the forthcoming century,” Proc. Nat. Acad. Sci. USA, vol. 107, no. 45, pp. 19167–19170, Nov. 2010.
[3] S. Hantson, S. Pueyo, and E. Chuvieco, “Global fire size distribution is driven by human impact and climate,” Global Ecology Biogeography, vol. 24, no. 1, pp. 77–86, Jan. 2015.
[4] S. Hantson, G. Lasslop, S. Kloster, and E. Chuvieco, “Anthropogenic effects on global mean fire size,” Int. J. Wildland Fire, vol. 24, no. 5, pp. 589–596, Aug. 2015.
[5] S. Archibald and D. P. Roy, “Identifying individual fires from satellite-derived burned area data,” Int. Geosci. Remote Sens. Symp. (IGARSS), vol. 3, Jul. 2009, pp. 160–163.
[6] S. Archibald, C. E. R. Lehmann, J. L. Gómez-Dans, and R. A. Bradstock, “Defining pyromes and global syndromes of fire regimes,” Proc. Nat. Acad. Sci. USA, vol. 110, no. 16, pp. 6442–6447, Apr. 2013.
[7] D. Oom, P. C. Silva, I. Bistinas, and J. M. C. Pereira, “Highlighting biome-specific sensitivity of fire size distributions to time-gap parameter using a new algorithm for fire event individuation,” Remote Sens., vol. 8, no. 8, p. 663, Aug. 2016.
[8] N. Andela et al., “The global fire atlas of individual fire size, duration, speed and direction,” Earth Syst. Sci. Data, vol. 11, no. 2, pp. 529–552, Apr. 2019.
[9] R. Whelan, The Ecology of Fire. Cambridge, U.K.: Cambridge Univ. Press, 1995.
[10] M. Fuller, Forest Fires: An Introduction to Wildland Fire Behavior, Management, Fighting, and Prevention. Hoboken, NJ, USA: Wiley, 1991.
[11] A. A. Ager, N. M. Vaillant, and M. A. Finney, “A comparison of landscape fuel treatment strategies to mitigate wildland fire risk in the urban interface and preserve old forest structure,” Forest Ecology Manage., vol. 259, no. 8, pp. 1556–1570, Mar. 2010.
[12] F. Moreira et al., “Landscape—Wildfire interactions in southern Europe: Implications for landscape management,” J. Environ. Manage., vol. 92, no. 10, pp. 2389–2402, Oct. 2011.
[13] J. M. P. Nogueira, J. Ruffault, E. Chuvieco, and F. Mouillot, “Can we go beyond burned area in the assessment of global remote sensing products with fire patch metrics,” Remote Sens., vol. 9, no. 1, p. 7, Dec. 2017.
[14] D. Frantz, M. Stellines, A. Röder, and J. Hill, “Fire spread from MODIS burned area data: Obtaining fire dynamics information for every single fire,” Int. J. Wildland Fire, vol. 25, no. 12, pp. 1228–1237, Jan. 2016.
[15] D. P. Roy, Y. Jin, P. E. Lewis, and C. O. Justice, “Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data,” Remote Sens. Environ., vol. 110, no. 16, pp. 6442–6447, Apr. 2015.
[16] L. Giglio, L. Boschetti, D. P. Roy, M. L. Humber, and C. O. Justice, “The Collection 6 MODIS burned area mapping algorithm and product,” Remote Sens. Environ., vol. 217, pp. 72–85, Nov. 2018.
[17] K. Tansey et al., “A new, global, multi-annual (2000–2007) burnt area product at 1 km resolution,” Geophys. Res. Lett., vol. 35, no. 1, Jan. 2008, Art. no. L01401.
[18] I. Alonso-Canas and E. Chuvieco, “Global burned area mapping from ENVISAT-MERIS and MODIS active fire data,” Remote Sens. Environ., vol. 163, pp. 140–152, Jan. 2015.
[19] F. Mouillot et al., “Ten years of global burned area products from spaceborne remote sensing—A review: Analysis of user needs and recommendations for future developments,” Int. J. Appl. Earth Observ. Geoinf., vol. 26, pp. 64–79, Feb. 2014.
[20] W. Seiler and P. J. Crutzen, “Estimates of gross and net fluxes of carbon between the biosphere and the atmosphere from biomass burning,” Clim. Change, vol. 2, no. 3, pp. 207–247, Sep. 1980.
