



## Mapping burned areas using dense time-series of Landsat data



Todd J. Hawbaker<sup>a,\*</sup>, Melanie K. Vanderhoof<sup>a</sup>, Yen-Ju Beal<sup>a</sup>, Joshua D. Takacs<sup>a</sup>, Gail L. Schmidt<sup>b</sup>, Jeff T. Falgout<sup>c</sup>, Brad Williams<sup>c</sup>, Nicole M. Fairaux<sup>a</sup>, Megan K. Caldwell<sup>d</sup>, Joshua J. Picotte<sup>e</sup>, Stephen M. Howard<sup>f,1</sup>, Susan Stitt<sup>a,1</sup>, John L. Dwyer<sup>f</sup>

<sup>a</sup> U.S. Geological Survey, Geosciences and Environmental Change Science Center, PO Box 25046, MS 980, Lakewood, CO 80225, USA

<sup>b</sup> Stinger Ghaffarian Technologies, Contractor to the U.S. Geological Survey, Earth Resources Observation and Science Center, 47914 252nd Street, Sioux Falls, SD 57198, USA

<sup>c</sup> U.S. Geological Survey, Core Science, Analytics, Synthesis, and Libraries, PO Box 25046, MS 302, Lakewood, CO 80225, USA

<sup>d</sup> Department of Ecology & Evolutionary Biology, University of Colorado Boulder, Boulder, CO 80309-0334, USA

<sup>e</sup> ASRC Federal InuTeq LLC, Contractor to the U.S. Geological Survey, Earth Resources Observation and Science Center, 47914 252nd Street, Sioux Falls, SD 57198, USA

<sup>f</sup> U.S. Geological Survey, Earth Resources Observation and Science Center, 47914 252nd Street, Sioux Falls, SD 57198, USA

### ARTICLE INFO

#### Article history:

Received 2 November 2016

Received in revised form 8 June 2017

Accepted 25 June 2017

Available online 7 July 2017

#### Keywords:

Burned area

Essential climate variable

Landsat

Wildfire

### ABSTRACT

Complete and accurate burned area data are needed to document patterns of fires, to quantify relationships between the patterns and drivers of fire occurrence, and to assess the impacts of fires on human and natural systems. Unfortunately, in many areas existing fire occurrence datasets are known to be incomplete. Consequently, the need to systematically collect burned area information has been recognized by the United Nations Framework Convention on Climate Change and the Intergovernmental Panel on Climate Change, which have both called for the production of essential climate variables (ECVs), including information about burned area. In this paper, we present an algorithm that identifies burned areas in dense time-series of Landsat data to produce the Landsat Burned Area Essential Climate Variable (BAECV) products. The algorithm uses gradient boosted regression models to generate burn probability surfaces using band values and spectral indices from individual Landsat scenes, lagged reference conditions, and change metrics between the scene and reference predictors. Burn classifications are generated from the burn probability surfaces using pixel-level thresholding in combination with a region growing process. The algorithm can be applied anywhere Landsat and training data are available. For this study, BAECV products were generated for the conterminous United States from 1984 through 2015. These products consist of pixel-level burn probabilities for each Landsat scene, in addition to, annual composites including: the maximum burn probability and a burn classification. We compared the BAECV burn classification products to the existing Global Fire Emissions Database (GFED; 1997–2015) and Monitoring Trends in Burn Severity (MTBS; 1984–2013) data. We found that the BAECV products mapped 36% more burned area than the GFED and 116% more burned area than MTBS. Differences between the BAECV products and the GFED were especially high in the West and East where the BAECV products mapped 32% and 88% more burned area, respectively. However, the BAECV products found less burned area than the GFED in regions with frequent agricultural fires. Compared to the MTBS data, the BAECV products identified 31% more burned area in the West, 312% more in the Great Plains, and 233% more in the East. Most pixels in the MTBS data were detected by the BAECV, regardless of burn severity. The BAECV products document patterns of fire similar to those in the GFED but also showed patterns of fire that are not well characterized by the existing MTBS data. We anticipate the BAECV products will be useful to studies that seek to understand past patterns of fire occurrence, the drivers that created them, and the impacts fires have on natural and human systems.

Published by Elsevier Inc.

### 1. Introduction

Accurate and complete data on fire locations and burned areas (prescribed and wild) are needed for a variety of applications including

quantifying trends and patterns of fire occurrence (Abatzoglou and Williams, 2016; Dennison et al., 2014; Giglio et al., 2013; Westerling et al., 2006); characterizing drivers of past fire occurrence and projecting future potential patterns of fires (Bachelet et al., 2003; Hawbaker et al., 2013; Krawchuk et al., 2009; Parisien and Moritz, 2009; Riley et al., 2013); and assessing the impacts of fires on a range of natural and social systems (French et al., 2014; Shakesby and Doerr, 2006; van der Werf et al., 2010; Williams et al., 2016). Many of these applications require

\* Corresponding author.

E-mail address: [tjhawbaker@usgs.gov](mailto:tjhawbaker@usgs.gov) (T.J. Hawbaker).

<sup>1</sup> Retired.

consistent fire data collected over long time periods to determine if changes in fire occurrence and fire impacts are related to shifts in climate, land-use/land-cover change, policy and management, and other drivers.

Recognizing the importance of fires, especially for understanding climate change and its impacts, the Global Climate Observing System (GCOS) included fire disturbance, specifically burned area, in their list of 13 terrestrial essential climate variables (ECVs) that are technically and economically feasible for systematic observation (Food and Agriculture Organization of the United Nations, 2008). These are being developed in response to calls from the United Nations Framework Convention on Climate Change, Intergovernmental Panel on Climate Change (Global Climate Observing System, 2004), as well as calls from the European Space Agency's Climate Change Initiative, the National Research Council, and the Landsat Science Team to systematically observe atmosphere, ocean, and land characteristics (Hollmann et al., 2013; National Research Council, 2001; Roy et al., 2014; Wulder et al., 2012).

Remote sensing is critical to the development of ECVs because satellite images and the products derived from them can provide a basis for long-term systematic data collection to monitor changes in the land surface that either influence or are influenced by climate. Monitoring active fires and burned areas, in particular, is feasible with remote sensing because of the thermal and spectral changes induced by fires. Coarse-resolution sensors equipped with spectral and thermal bands, such as the Advanced Very High Resolution Radiometer (AVHRR), can effectively map subpixel heat sources (Dozier, 1981; Matson and Dozier, 1981) and applications for spatially-extensive active fire detection with AVHRR data were realized (Flannigan and Vonder Haar, 1986; Malingreau et al., 1985; Matson et al., 1987) and extended to other sensors such as the Geostationary Operational Environmental Satellite Visible Infrared Spin Scan Radiometer Atmospheric Sounder (Prins and Menzel, 1992; Prins and Menzel, 1994), the Defense Meteorological Satellite Program Operational Linescan System (Elvidge et al., 1996), the Moderate Resolution Imaging Spectroradiometer (MODIS) (Giglio et al., 2003), the Visible Infrared Imaging Radiometer Suite (VIIRS) (Giglio et al., 2000), and even moderate-resolution sensors like the Landsat Thermal Infrared Sensor (Schroeder et al., 2015). In addition to identifying actively burning fires, spectral changes visible in the temporally rich data provided by coarse-resolution sensors allowed for burned area detection, although the approaches varied depending on the spectral bands specific to individual sensors (Eva and Lambin, 1998; Kasischke and French, 1995; Roy et al., 2005). Other approaches combined spectral change analysis with hot spot detection, to help distinguish burned areas from other types of change (Alonso-Canas and Chuvieco, 2015; Fraser et al., 2000; Giglio et al., 2009; Li et al., 1997). Many of the existing coarse resolution global active fire and burned area algorithms and products provide the types of data identified in the GCOS ECV definitions (Alonso-Canas and Chuvieco, 2015; Global Climate Observing System, 2004) and have been combined to produce operational products monitoring burned areas and emissions such as the Global Fire Emissions Database (GFED; Giglio et al., 2013; Randerson et al., 2012; van der Werf et al., 2010).

In spite of the impressive efforts made to monitor fire activity and burned area with coarse-resolution sensors, many shortcomings remain. Fires obscured by clouds, fires with short-lived thermal signatures, and small fires may not be identified (Hawbaker et al., 2008; Morisette et al., 2005; Schroeder et al., 2008). Detection errors remain high for many global coarse-resolution products; Padilla et al. (2015) found that commission errors for burned area ranged between 42% and 94%, and omission errors for burned area ranged between 68% and 93%, depending on the sensor used. Additionally, the short time series provided by coarse-resolution sensors (other than AVHRR) are inadequate when trying to quantify relationships between patterns of climate and fire occurrence because they span a limited range of climate variability, potentially over emphasizing the importance of extreme years (Hawbaker et al., 2013; Hawbaker and Zhu, 2012; Westerling et al., 2011).

Moderate-resolution sensors such as the Landsat Multispectral Scanner System (MSS) and Thematic Mapper (TM) have also been used to remotely sense burned areas. Initial efforts used Landsat pre- and post-fire images to map pre-fire vegetation, burned area extent, and severity (Chuvieco and Congalton, 1988; Hall et al., 1980; Jakubauskas et al., 1990; Koutsias and Karteris, 1998). Subsequent efforts focused on identifying burned areas within single scenes (Koutsias and Karteris, 2000; Kushla and Ripple, 1998), extracting within-fire heterogeneity, severity, mortality, and carbon loss (Michalek et al., 2000; Miller and Yool, 2002; Patterson and Yool, 1998; Rogan and Yool, 2001). The moderate resolution of Landsat sensors also allowed for the development of burned area detection algorithms using spatial contagion metrics and region-growing approaches to incorporate the spatial patterns of spectral reflectance among neighboring pixels, in addition to the pixel-level spectral data to identify burned areas; helping to reduce omission errors (Bastarrika et al., 2011; Chuvieco et al., 2002; Goodwin and Collett, 2014; Koutsias, 2003; Stroppiana et al., 2012). The results of these approaches provide data products with the spatial and temporal resolution relevant to fire ecology and management and have laid the foundation for operational programs monitoring patterns of severity and area burned by large fires in the United States (Eidenshink et al., 2007).

Since the opening of the Landsat archive (Wulder et al., 2012), change detection approaches for moderate-resolution data have evolved to incorporate the temporal depth of the data available in the Landsat archive by subdividing annual time series of spectral responses into piecewise segments, and then using the changes between segments and characteristics of segments to delineate disturbances, such as the Vegetation Change Tracker (VCT; Huang et al., 2010) and the Landsat-based detection of Trends in Disturbance and Recovery (LandTrendr; Kennedy et al., 2010) algorithms. More recent approaches have analyzed the time series as a whole (Hansen et al., 2014) or decomposed dense time series data to distinguish seasonality from long-term trends for change detection using both MODIS (Verbesselt et al., 2010) and Landsat data (Brooks et al., 2014; Zhu and Woodcock, 2014b). These methods detect change, but require additional attribution to characterize the specific type of change (Kennedy et al., 2015; Liang et al., 2014; Zhao et al., 2015). The methods cited above have largely focused on detection of stand-replacing forest disturbances, or in the case of fires, those that result in long-lasting changes in spectral reflectance visible in annual Landsat time series stacks (e.g. stand-replacing forest fires, Huang et al., 2010; Kennedy et al., 2010).

Fire-specific change detection algorithms making use of the full temporal depth of the Landsat archive are also emerging. For example, Goodwin and Collett (2014) combined a change-detection algorithm with a region-growing algorithm to identify contiguous areas of change and then classified which areas of change were caused by fires in Queensland Australia, using all available Landsat data from 1986 through 2013. Similarly, Boschetti et al. (2015) used weekly composites of Landsat 7 data from 2002 to identify areas with spectral change and then combined them with MODIS active fire data to separate burned areas in predominantly forested regions in the western U.S. These two studies demonstrated that burned areas can be mapped with automated approaches over large spatial extents using moderate-resolution data. However, the need for consistently collected burned area data covering large spatial (national, continental, and global scales) and temporal (30 or more years) extents has not been met.

In this paper, we present an algorithm to identify burned areas in Landsat satellite images, compare its outputs with existing burned area dataset, and discuss how those outputs or products provide novel information about patterns of fire occurrence in the conterminous United States (CONUS). The first objective of this study was to develop a new algorithm to identify burned areas in satellite imagery; henceforth referred to as the Landsat Burned Area Essential Climate Variable (BAECV) algorithm. The second objective of this study was to produce a publicly available Landsat-based burned area product from 1984

through 2015 for the CONUS. To do this, we applied the BAECV algorithm to the CONUS. To demonstrate that the BAECV products provide new information, we compared them with the GFED and MTBS burned area data.

## 2. Methods

### 2.1. Study area and Landsat scene selection

We selected the CONUS as our study area because it has a large amount of both Landsat data and fire occurrence data available. For any given area, a single Landsat satellite can acquire 23 observations per year; with two satellites operating simultaneously this increases to 46 observations per year. To train and evaluate our approach, we selected 38 Landsat World-Reference System 2 (WRS-2) path and rows (path/rows) across the CONUS (Fig. 1). Path/row locations were spatially distributed in order to capture major ecosystems and fire regimes, largely based on the U.S. Environmental Protection Agency's Level 2 and 3 Ecoregions (Omernik and Griffith, 2014). However, we also included scenes where we expected burned area detection with Landsat data to be especially challenging, such as path and rows in the Great Plains and semi-arid regions of the Southwest.

We developed the BAECV algorithm specifically to operate on Landsat imagery, since these data span a long time series (1972–present, depending on the sensor), are freely available, and have moderate spatial resolution (Wulder et al., 2012). Datasets used in this analysis included Landsat 4 Thematic Mapper (TM; 1982–1993), Landsat 5 TM (1984–2013), Landsat 7 Enhanced Thematic Mapper Plus (ETM+) both with the Scan Line Corrector (SLC) on (1999–2003) and the SLC off (2003–2015). Using the U.S. Geological Survey EarthExplorer website ([earthexplorer.usgs.gov](http://earthexplorer.usgs.gov)), we selected all available scenes with (1) precision and terrain correction (Level 1T), (2) cloud cover less than or equal to 80%, and (3) georeferencing root mean square error  $\leq 10$  m. The scenes found in EarthExplorer were then submitted to the Earth Resources Observation and Science (EROS) Center Science Processing Architecture (ESPA) Ordering Interface ([espa.cr.usgs.gov](http://espa.cr.usgs.gov)) for bulk processing. The Landsat Ecosystem Disturbance Adaptive Processing System generated surface reflectance (Masek et al., 2006) and the source metadata were ordered through ESPA. This resulted in 22,855 Landsat scenes for training and evaluation.

### 2.2. Landsat Burned Area Essential Climate Variable algorithm

The BAECV algorithm is a supervised approach that uses gradient boosted regression models to estimate the probability that each pixel within a Landsat image was burned. The burn probabilities are then composited to per-pixel maximum annual burn probabilities, followed by a thresholding and segmentation process to generate binary annual

burn classification images. Before describing the algorithm, we provide details about the predictor and response variables derived from Landsat data and other sources, and the selection of sample points used for training and testing the algorithm.

#### 2.2.1. Fire occurrence data

The MTBS data (Eidenshink et al., 2007) were the primary source of observed burned areas used for training and evaluating the results of the BAECV algorithm. These data include large fires ( $\geq 2$  km<sup>2</sup> in the eastern U.S. and  $\geq 4$  km<sup>2</sup> in the western U.S.). We selected the MTBS dataset for use in this analysis because it is the most comprehensive spatially-explicit burned area dataset for the U.S. Even though MTBS does not map all fires, each fire in the dataset has been visually interpreted by a trained analyst. Data for each MTBS fire consist of a fire perimeter polygon and a categorical burn severity raster layer derived from either a single post-fire Landsat image or a pair of pre- and post-fire Landsat images. The MTBS fires used in this study span the Landsat 4, 5, and 7 epochs (1984–2013); and included approximately 17,637 fires that burned  $>445,000$  km<sup>2</sup> across the CONUS. For this analysis, the MTBS fire perimeter polygons and severity rasters were cropped to the 38 training path/row polygons before additional processing; resulting in our training and testing data containing 4493 individual fires burning 142,000 km<sup>2</sup>.

#### 2.2.2. Predictor variables

The BAECV algorithm makes use of three types of predictor variables: (1) land surface conditions in the *scene* or image of interest; (2) *reference* conditions that characterize the land surface prior to disturbances visible in the scene of interest; and (3) *change* metrics quantifying the magnitude of change between the reference conditions and the scene (Table 1). Scene-level predictors used by the BAECV algorithm included the individual Landsat bands in addition to a number of spectral indices. The reference variables represented pre-fire surface conditions using three-year lagged means and standard deviations for each of the single-scene predictor variables. The change predictor variables included absolute and relative difference between the values of a given predictor in a Landsat scene (e.g. NBR) and its corresponding reference predictor (e.g. the 3-year lagged mean of NBR). Additional predictors included the Landsat sensor number (e.g. 4, 5, or 7) and the U.S. Environmental Protection Agency's Level I ecoregions (Omernik and Griffith, 2014). Pixels flagged as cloudy, water, or snow/ice by the function of mask algorithm (FMask) (Zhu and Woodcock, 2014a) were excluded from calculations.

#### 2.2.3. Sampling points for training and testing

To generate a sample of training and testing data points, we randomly selected points within the polygon outlining the extent of each WRS-2 path and row (Fig. 1). To sample burned areas, we randomly selected

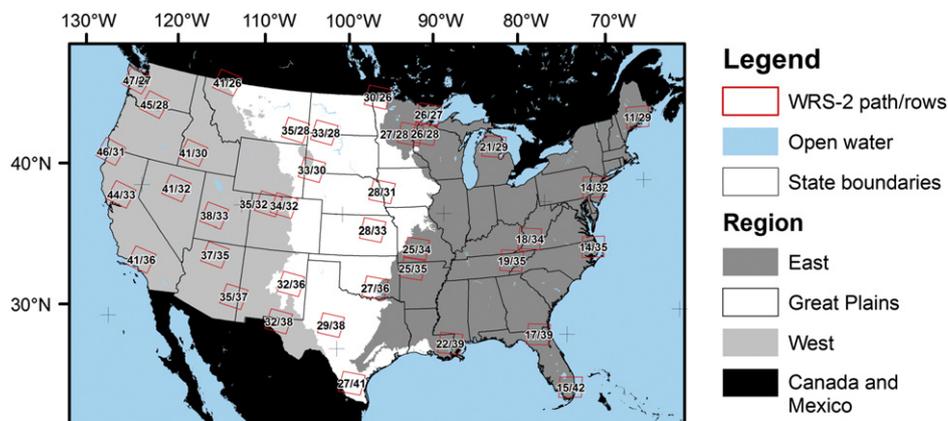


Fig. 1. Study area showing the World Reference System version 2 (WRS-2) path/rows where training data were collected for the Landsat Burned Area Essential Climate Variable algorithm.

**Table 1**

Landsat Thematic Mapper and Enhanced Thematic Mapper Plus derived predictors used in the gradient boosted regression models. NIR: near infrared; SMIR: shortwave middle infrared; LMIR: longwave middle infrared.

Variable	Abbreviation	Formula	Reference
Blue: 0.45–0.52 $\mu\text{m}$	band1		
Green: 0.52–0.60 $\mu\text{m}$	band2		
Red: 0.63–0.69 $\mu\text{m}$	band3		
NIR: 0.76–0.90 $\mu\text{m}^a$	band4		
SMIR: 1.55–1.75 $\mu\text{m}$	band5		
Thermal: 10.40–12.50 $\mu\text{m}$	band6		
LMIR: 2.08–2.35 $\mu\text{m}^b$	band7		
Normalized difference vegetation index	NDVI	$(\text{band4} - \text{band3}) / (\text{band4} + \text{band3})$	Tucker, 1979
Normalized difference moisture index	NDMI	$(\text{band4} - \text{band5}) / (\text{band4} + \text{band5})$	Gao, 1996
Normalized difference wetness index	NDWI	$(\text{band2} - \text{band4}) / (\text{band2} + \text{band4})$	McFeeters, 1996
Tasseled cap greenness	TC greenness	$(-0.1603 * \text{band1}) + (-0.2819 * \text{band2}) + (-0.4934 * \text{band3}) + (0.7940 * \text{band4}) + (0.0002 * \text{band5}) + (-0.1446 * \text{band7})$	Crist, 1985
Tasseled cap wetness	TC wetness	$(0.0315 * \text{band1}) + (0.2021 * \text{band2}) + (0.3102 * \text{band3}) + (0.1594 * \text{band4}) + (0.6806 * \text{band5}) + (-0.6109 * \text{band7})$	
Normalized burn ratio	NBR	$(\text{band4} - \text{band7}) / (\text{band4} + \text{band7})$	López García and Caselles, 1991;
Normalized burn ratio 2	NRB2	$(\text{band5} - \text{band7}) / (\text{band5} + \text{band7})$	Key and Benson, 2006

3-Year lagged mean and standard deviations for each spectral index.

Absolute and relative change from 3-year lagged mean for each spectral index.

<sup>a</sup> 0.77–0.90  $\mu\text{m}$  for ETM+.

<sup>b</sup> 2.09–2.35  $\mu\text{m}$  for ETM+.

up to 2000 point locations within MTBS perimeters in each path and row. For this analysis, we only considered MTBS points classified as low, moderate, or high burn severity to be burned areas. The date of the fire was defined by the MTBS perimeters and was assigned to each point. To sample unburned areas, we also randomly selected an equal number of point locations outside of the MTBS perimeters. We specified a minimum distance of 30 m between points to ensure that individual pixel locations were only sampled once. At each point location, and for each image in the time series, we collected the values for the ecoregion, sensor id (4, 5, or 7), and each of the single-scene predictors, 3-year lagged predictors, and absolute and relative change between the single-scene predictors and the 3-year lagged predictors (Table 1). These data were split into training and testing groups based on years. Six years were retained for testing and validation (1988, 1993, 1998, 2003, 2008, and 2013). The remaining 24 years of data were used for training. After the training and testing split, all points labeled as burned from the MTBS data were retained and an equally-sized sample of unburned points was randomly selected.

#### 2.2.4. Burned area probability mapping

The first step of the BAECV algorithm is to estimate the probability that each pixel in a Landsat image had burned. To do this, we first fit, and later made predictions with a gradient boosted regression model (GBRM; Hastie et al., 2009) which uses a sequence of simple classification and regression tree (CART) models (Breiman et al., 1984). Classification and regression trees have been used extensively in remote sensing (Chan et al., 2001; Friedl and Brodley, 1997; Rogan et al., 2002) because they have many advantages over other types of statistical models, including: the ability to handle both categorical and continuous variables; allow for missing data; include variable interactions in models with more than one split; make no assumptions about the model structure or distribution of predictor variables; rarely select unimportant variables; do not degrade predictive accuracy when highly correlated predictors are included; and produce easily interpreted results (Hastie et al., 2009). However, CART models also have a number of disadvantages including: instability, or small changes in the input data resulting in large changes in splitting thresholds; smooth functions or relationships between predictor and response variables may not be well characterized; and over-fitting may occur when large trees are not pruned. The shortcomings over CART models have been circumvented by using ensemble methods such as bagging and boosting (Hastie et al., 2009). The bagging approach is used by random forests which fit many individual CART models, sampling a subset of the data

for each model, and averaging their results – often out performing individual CART models. The boosting approach, as implemented by GBMRMs, is similar to random forests; however, the CART models are fit in a sequence. There is also a learning component; the observations used to fit each CART model are weighted by the residuals of the previous CART model. The resulting GBRM is a committee of learners that often has higher predictive accuracy than that of individual CART, and random forest models for binary classification problems (Hastie et al., 2009). All 3 types of modeling approaches: CART, random forests, and GBMRMs have been used for a wide range of remote sensing applications (Coulston et al., 2012; Hansen et al., 2014; Homer et al., 2004; Liu, 2016; Merentitis and Debes, 2015; Powell et al., 2010; Schneider, 2012; Thompson et al., 2016).

Training a GBRM requires specification of a number of parameters that control the model structure. These parameters include the (1) number of trees, (2) number of splits per tree, and (3) learning rate between successive trees (Hastie et al., 2009). Individual trees are fit in sequence and when fitting the tree, the learning rate specifies the weight to apply to prediction errors from the previous tree in the sequence. In practice, setting these three parameters involves fitting models for a range of values and then evaluating tradeoffs among model complexity, accuracy of predictions, and computation time.

We selected a range of values for the learning rate (0.1, 0.05, and 0.01) and the number of splits per tree (1, 3, and 5). For each combination of learning rate and number of splits per tree, we fit GBMRMs with our training data using 5000 trees. After initial model fits, the number of trees used in each of the GBMRMs was systematically reduced by evaluating changes in the loss metric for the test data as a function of the tree's number in the sequence trees. The objective here was to determine the smallest number of trees in the sequence needed to achieve the maximum value of an accuracy metric. To compare predictive accuracy among the final GBMRMs, we calculated the area under the curve (AUC) of receiver-operating characteristic plots metric using the test data. After fitting the GBMRMs, we selected the individual GBRM that had the highest accuracy with the lowest number of trees (to reduce computation time). The final GBRM was used to generate pixel-level burn probability images for all images in the entire time-series of Landsat scenes, which were subsequently used to generate annual composites of maximum burn probability.

#### 2.2.5. Burned area classification

The next step in our approach was to threshold the annual composites of maximum burn probability to produce burn classifications, or

binary images specifying which pixels were burned and not burned. Through visual analysis, we observed that burn probability images and annual composites often had patches of pixels with very-high burn probabilities, and that those patches were often connected by pixels with slightly lower, but still high burn probabilities within MTBS fire perimeters and much lower burn probabilities outside of the MTBS fire perimeters. To capture this observed pattern, we implemented a region-growing method similar to approaches taken in other studies (Bastarrika et al., 2011; Chuvieco et al., 2002; Goodwin and Collett, 2014; Koutsias, 2003; Stroppiana et al., 2012). We first identified seeds for potential burned area regions by thresholding the annual maximum burn probability composites into preliminary binary burned patches; this threshold value is referred to as the 'seed probability threshold'. Next, burned area patches below a minimum size were removed, specified using the 'seed size threshold'. Third, neighboring pixels with burn probabilities slightly lower than the seed probability threshold were added to patches if the neighboring pixels had a burn probability greater than or equal to the 'spread probability threshold'. The third step was completed in an iterative fashion until no additional neighboring pixels with burn probabilities above the spread probability threshold could be found. The threshold values we used were 98% for the seed probability threshold, 45 pixels for the seed size threshold, and 95% for the spread probability threshold. Once the three thresholds were set, an annual burned area classification product was generated for each year from 1984 to 2015.

### 2.3. Evaluation of algorithm and outputs

In this study, we evaluated the BAECV algorithm and its resulting products in a number of ways. To assess the algorithm performance, we examined the relative importance and partial dependence plots of predictor variables to identify which variables were contributing the most to the BAECV models and how they related to the probability that a pixel had burned. Additionally, we examined a number of output images to assess how well the BAECV algorithm extracts burned areas in time-series of Landsat images and to provide illustrative examples of where the algorithm works well and doesn't work so well.

To evaluate the BAECV products, we summarized the percent of 1/4-degree grid cells burned using the BAECV data, and compared those results to equivalent summaries from the GFED (version 4.1s) and MTBS data for CONUS and 3 major regions of the CONUS: West, Great Plains, and East (Fig. 1).

We made more detailed comparisons between the BAECV and MTBS data because they had the same spatial resolution and nearly same temporal extent. Specifically, we assessed differences in burned area between BAECV and MTBS products in relation to land cover from 1993 to 2013 (land cover data prior to 1992 were not available). In addition to quantifying differences between the BAECV and MTBS products across land cover types, we also assessed what percentage of MTBS pixels were not mapped by the BAECV algorithm by MTBS burn severity class (low, moderate, or high). Finally, we used the BAECV and MTBS data to compare rates of change in burned area between 10-year increments (1984–1993; 1994–2003; and 2004–2013).

Data preparation prior to analysis involved combining the BAECV burn classification mosaics, MTBS annual burn severity mosaics, land cover from the National Land Cover Database (NLCD) (Fry et al., 2011; Homer et al., 2007; Homer et al., 2015; Vogelman et al., 2001), and a raster delineating the three regions of the CONUS. Because NLCD classifications vary across the different versions, we reassigned them using a consistent classification scheme (Table 2) prior to analysis. Land cover classes that are unable or unlikely to burn: open water, perennial ice/snow, developed, and barren areas were removed. The version of NLCD data used in the combined dataset was assigned based on year of the burned area data. For example, the 1992 NLCD data were combined with burned area data from 1992 through 2000.

## 3. Results

### 3.1. Algorithm structure and performance

After fitting all possible combinations of GBRM parameters (number of splits/tree and learning rate), we selected 5 splits/tree and a 0.1 learning rate as the 'best' combination because it had the lowest number of trees required to reach the maximum value of the accuracy metric, calculated from the test sample (AUC; Tables 3 and 4). Possible AUC values range between 0.0 and 1.0 and represents the probability that the model is able to correctly identify a burned pixel, given a pair of pixels (1 burned and 1 unburned). Values of 0.5 indicate the model is no better than a random guess. Model accuracy improves as AUC approaches a value of 1.0, which would indicate the model is a perfect classifier. Our GBRM had an AUC value of 0.89 suggesting that its classification accuracy was quite good.

We assessed the relative importance of the predictor variables used in the final GBRM. We found that the ecoregion, individual Landsat bands, and the 3-year lagged summary predictors had the greatest relative importance, while the scene-level spectral indices and change metrics were less important (Fig. 2). Partial dependence plots showed that burn probability had a negative relationship with the shortwave middle infrared (band5) and a positive relationship with the longwave middle infrared (band7; Fig. 3). These relationships were amplified by the NBR2 spectral index which had an abrupt drop in burn probability in the 0.1 to 0.2 range (Fig. 3). The partial dependence plots for the 3-year lagged mean and standard deviation of NBR2 showed burn probability had a positive relationship with the 3-year lagged mean of NBR2 and a negative relationship with the 3-year lagged standard deviation of NBR2. Finally, the difference between NBR2 and the 3-year lagged mean of NBR2 had a rapidly changing relationship to burn probability, similar to the NBR2 relationship (Fig. 3).

To determine if the BAECV algorithm was producing reasonable results, we visually compared the BAECV burn probability and classification images with the original Landsat imagery (Fig. 4). The example for California (Fig. 4a) is an area dominated by grassland and shrubland with some coniferous forest. It shows that SLC gaps in Landsat 7 data present some challenges, leaving gaps in the corresponding burn probability image. However, the SLC gaps are often removed in the annual burn classification (shown with black outlines) because every image in the Landsat time series is assessed by the BAECV algorithm.

Fig. 4b shows part of the Hinman Fire that occurred in 2002 in northern Colorado. Land cover in this area was primarily coniferous forest and shrubland. This example shows that the burn probability image clearly identifies the burned area, but unburned areas contained some noise. These were ultimately filtered out by the burn classification process.

Fig. 4c shows an example of burned area detection in grasslands in eastern Kansas. Spring prescribed fires are frequently used in this region to promote grass production for cattle grazing. The perennial grasses and forbs in this area recover quickly after fires and burned areas might be missed if cloud-free imagery is not available in the spring.

The example for western Wisconsin (Fig. 4d) is located in a diverse landscape with a mix of deciduous forest, forested wetlands, pasture/hay, and agriculture land cover types. A clearly visible burned area was identified from the image shown and additional burned area (delineated by the black line) was identified in later images. In the southwest corner of this image, a small wetland area was incorrectly identified as burned, possibly because of changes in surface water extent.

In the North Carolina example (Fig. 4e), an active fire is shown burning in forested wetlands in the Great Dismal Swamp National Wildlife Refuge in 2011. Some of the burned area had relatively low probabilities and this may be in part because the area shown had previously burned in 2008 and not fully recovered by 2011, or due to the presence of smoke plumes in some of the images. Within the burn probability images, some burned areas were incorrectly masked as open water by

**Table 2**

Values used to standardize the different vintages of the National Land Cover Database (NLCD) to a common set of values.

Original Value	NLCD 1992 description	NLCD 2001, 2006, and 2011 description	Standardized description	Standardized value
11	Open water	Open water	Open water	11
12	Perennial ice/snow	Perennial ice/snow	Perennial ice/snow	12
21	Developed, open space	Developed, open space	Developed	20
22	Developed, high intensity	Developed, low intensity	Developed	20
23	Commercial/industrial/transportation	Developed, medium intensity	Developed	20
24		Developed, high intensity	Developed	20
31	Bare rock/sand/clay	Barren	Barren	31
32	Quarries/strip mines/gravel pits		Barren	31
33	Transitional		Shrub/scrub	52
41	Deciduous forest	Deciduous forest	Forest	40
42	Evergreen forest	Evergreen forest	Forest	40
43	Mixed forest	Mixed forest	Forest	40
51	Shrubland		Shrub/scrub	52
52		Shrub/scrub	Shrub/scrub	52
61	Orchards/vineyards/other		Cultivated crops	82
71	Grassland/herbaceous	Grassland/herbaceous	Grassland/herbaceous	71
81	Pasture/hay	Pasture/hay	Pasture/hay	81
82	Cultivated crops	Cultivated crops	Cultivated crops	82
83	Small grains		Cultivated crops	82
84	Fallow		Pasture/hay	81
85	Urban/recreational grasses		Developed	20
90	Woody wetlands	Woody wetlands	Wetlands	90
91	Woody wetlands		Wetlands	90
92	Emergent herbaceous wetlands		Wetlands	90
95	Emergent herbaceous wetlands	Emergent herbaceous wetlands	Wetlands	90

the FMask algorithm. In all five cases, the BAECV burn probability images clearly delineate the burned and unburned areas well and provide confidence that the BAECV algorithm is capable of detecting burned areas across a wide range of ecosystem types.

An example of the Landsat time series and corresponding BAECV products in southern Florida is shown in Fig. 5. Numerous burned areas appear and disappear as biomass is consumed by fires and subsequently recovers over time. Some areas were more clearly delineated in the burn probability images than others (e.g. Julian dates 43 and 59 vs. 107 and 315). Burned areas were visible in images with some cloud cover (e.g. Julian date 75, 139, 187, and 315). Actively burning fires in agricultural areas are visible for Julian dates 299 and 315 and some burned areas were detected in the agricultural areas in the top of those images. The final burn classification for 1988 was most likely an underestimate of burned area as some areas were mapped with low burn probabilities but were not retained by the thresholding step in our algorithm.

### 3.2. Spatial patterns of burned area

We compared the amount of area burned in the BAECV products, the GFED (version 4.1s; 1997–2015), and MTBS (1984–2013) across the CONUS and three regions of CONUS: the West, Great Plains, and the East. Between 1997 and 2015, the BAECV products mapped 36% more burned area than the GFED (722,887 vs. 532,992 km<sup>2</sup>; Table 5). The percent of burned area in the BAECV products was nearly equal between the West (40% of CONUS total) and the East (37%), and somewhat lower in the Great Plains (23%). In contrast, burned area in the GFED was primarily in the West (41% of CONUS total), followed by the Great Plains (32%), and then the East (27%). Differences between the BAECV and GFED burned area were greatest in the East (BAECV mapped 88%

more than GFED; Table 5) and the West (BAECV mapped 32% more), and similar in the Great Plains (BAECV mapped 4% less). The spatial pattern of the differences between the BAECV and GFED are visible in ¼-degree gridded summaries (Fig. 6).

We also compared the BAECV products to the MTBS data (only data from 1984 to 2013 were used as MTBS data were not available for 2014 and 2015). From 1984 through 2013, a total of 936,057 km<sup>2</sup> of burned area was identified in the BAECV products (Fig. 7). The majority of the area burned was in the West (39%) and Great Plains (37%), and a smaller percentage in the East (24%). Compared to the MTBS data, the BAECV products mapped 116% more burned area (Table 6; Fig. 7). In the MTBS data, a greater percentage of the burned area was in the West (65%), compared to the Great Plains (19%) and the East (16%). Differences in annual burned area were evident between the two datasets (Table 6) and especially large in the Great Plains (BAECV mapped 312% more than MTBS) and the East (BAECV mapped 233% more than MTBS), and slightly lower but still relatively large in the West (BAECV mapped 31% more than MTBS). The spatial pattern of these differences is visible in ¼-degree gridded summaries of BAECV and MTBS burned area, and their differences (Fig. 7).

Across the CONUS, 69% of the burned area in the BAECV products was not present in the MTBS data (Table 7). The largest amount of 'new' burned area from the BAECV algorithm was in the East and the Great Plains (87% and 83%, respectively). The majority of the burned area in the MTBS data was mapped in the BAECV products (67% across CONUS), but with substantial regional variability. The BAECV products identified burned areas that were also in the MTBS products in the West and Great Plains well, 73% and 69%, respectively, but missed a substantial portion (43%) of the MTBS burned area in the East.

**Table 3**

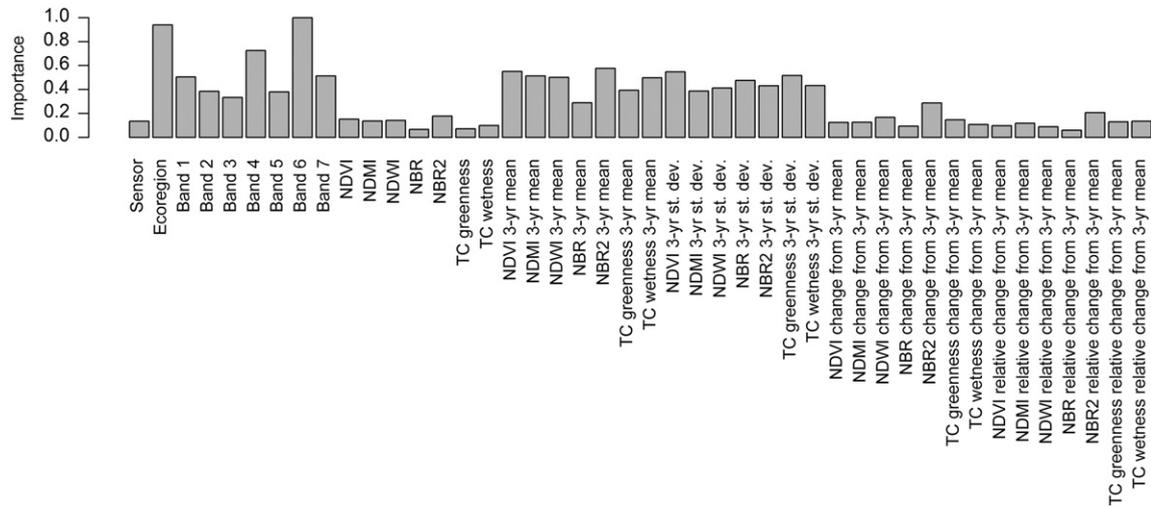
Number of trees for different gradient boosted regression model parameter combinations.

Number of splits/tree	Learning rate		
	0.1	0.05	0.01
1	2000	2500	2000
3	1500	1750	3000
5	750	1500	2250

**Table 4**

Area under the curve (AUC) of receiver-operating characteristic plots for different gradient boosted regression model parameter combinations.

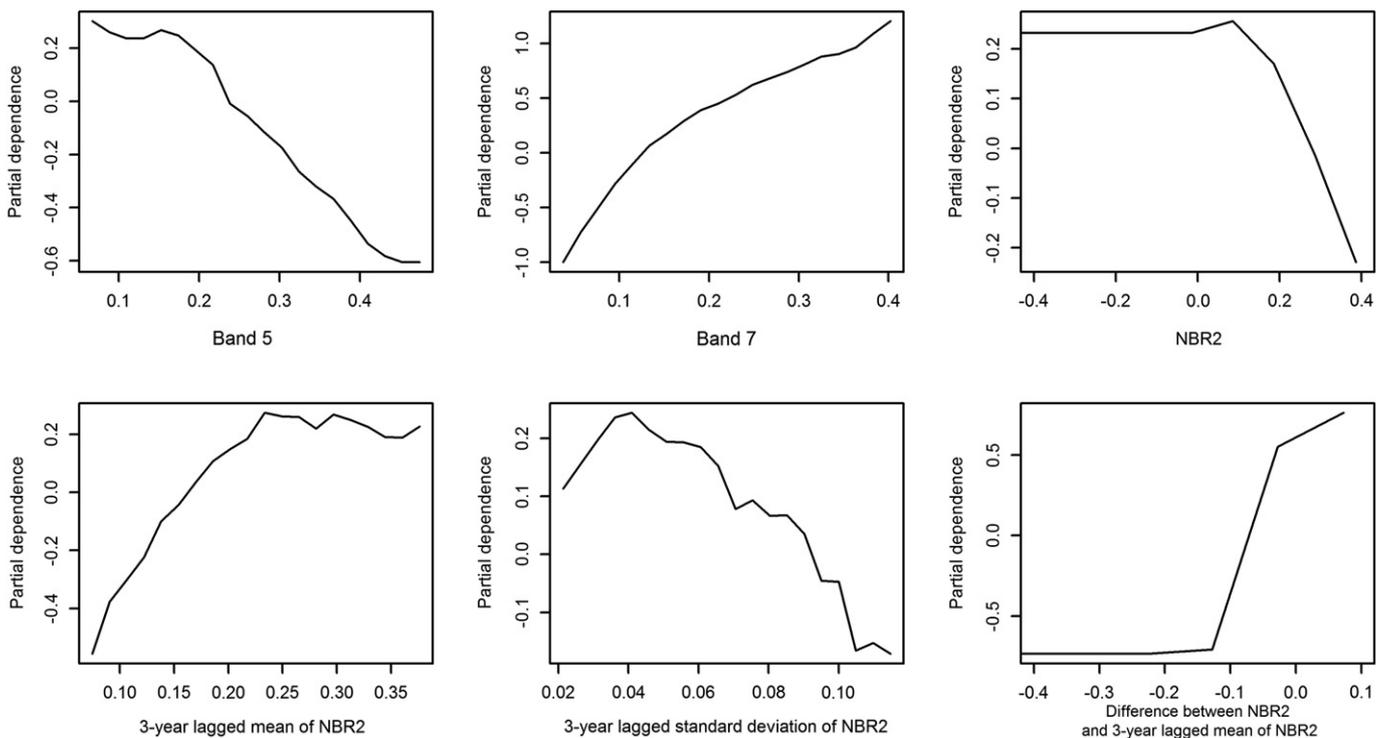
Number of splits/tree	Learning rate		
	0.1	0.05	0.01
1	0.83	0.82	0.77
3	0.88	0.88	0.84
5	0.89	0.89	0.89



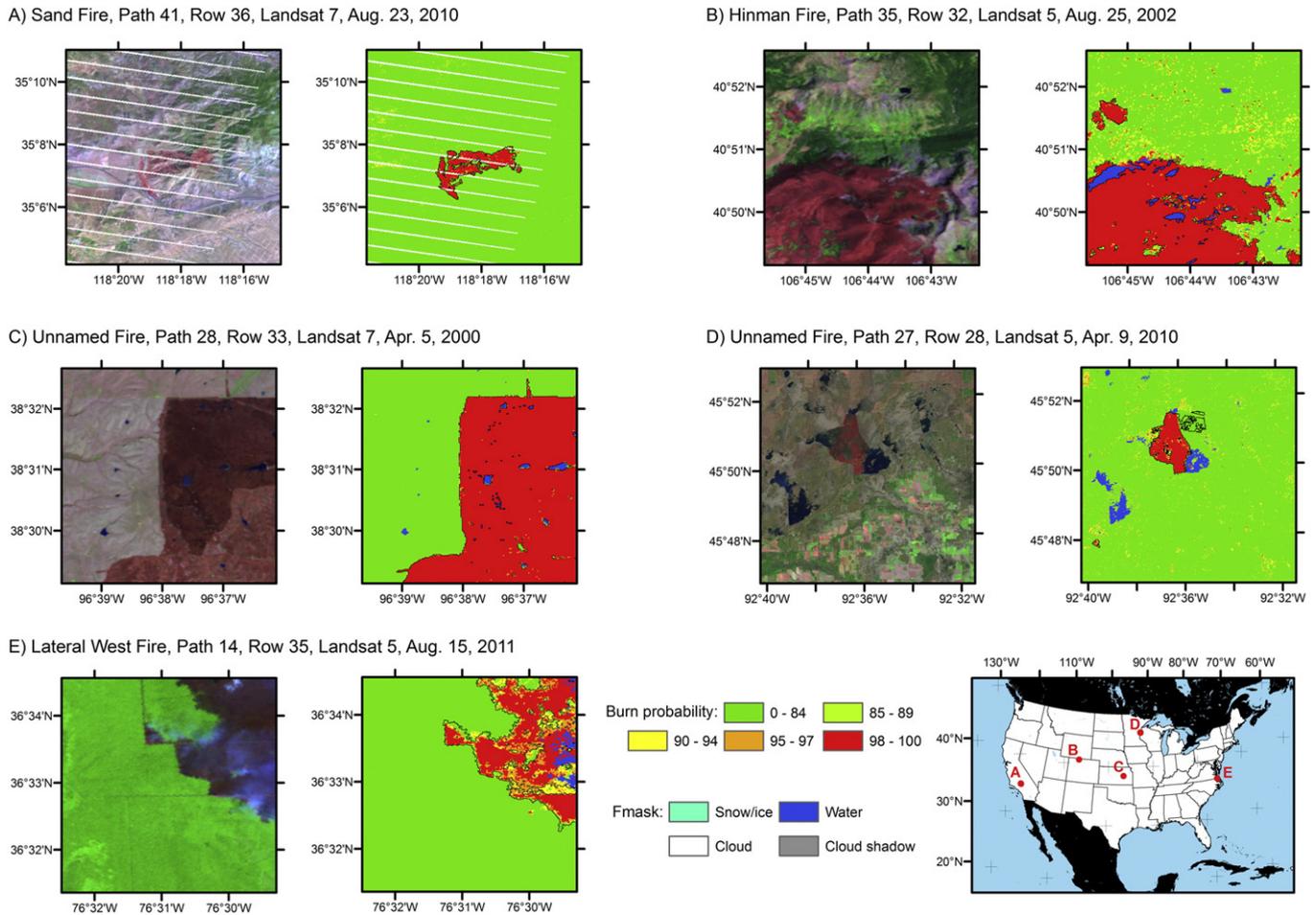
**Fig. 2.** Relative importance of individual predictor variables used in the gradient boosted regression model (GBRM). Results shown for GBRM using 5 splits/tree and a 0.1 learning rate. Bands 1–7: Landsat Thematic Mapper and Enhanced Thematic Mapper Plus bands; NDVI: normalized difference vegetation index; NDMI: normalized difference moisture index; NDWI: normalized difference wetness index; NBR: normalized burn ratio; NBR2: normalized burn ratio version 2; TC: tasseled cap; st. dev.: standardized deviation.

The majority of burned area in the BAECV products was in grassland/herbaceous (26%), shrub/scrub (25%), and forest (25%) land-cover classes. Smaller amounts of burned area were found in agriculture (10%), pasture/hay (7%), and wetland (6%) land-cover classes. These percentages were different for the MTBS data, which had a larger proportion of burned area in shrub/scrub (39%) and forest (31%) and a lesser proportion in grassland/herbaceous (21%), agriculture (1%), and pasture (1%), but a similar amount of wetland burned area (7%). In the West, burned area was spread across forest, shrub/scrub, and to some extent, the grassland/herbaceous land-cover classes while burned areas were primarily in the grassland/herbaceous and agriculture land-cover classes in the Great Plains, and primarily in the forest and wetland classes in the East (Fig. 8).

In addition to quantifying differences in burned area between the BAECV and MTBS products, we also evaluated potential differences in detection related to fire severity. The BAECV detected the majority of low (70%), moderate (87%) and high severity (85%) pixels mapped by MTBS across CONUS (Table 8). The BAECV performed weakest for low severity pixels in the East (46% detection rate); however, most fire regions showed detection rates of >70% across the three severity categories indicating that the BAECV maps most of the pixels mapped by the MTBS product (Table 8). A typical example of this comparison is shown in Fig. 9, in which the BAECV mapped both fires but was more conservative in defining the burned area extent relative to the reference dataset and the MTBS mapped fire severity.



**Fig. 3.** Partial dependence plots for a subset of predictor variables used in the gradient boosted regression model. NBR2: normalized burn ratio version 2.



**Fig. 4.** Example Landsat images and Burned Area Essential Climate Variable burn probability images from (A) California, (B) Colorado, (C) Kansas, (D) Wisconsin, and (E) North Carolina. Areas that were classified as burned are outlined in black.

### 3.3. Temporal trends of burned area

The BAECV annual burn classifications show distinct patterns of burned area and those patterns varied over time (Fig. 10; Fig. 11). For instance, in the 1984–1993 time period, large fires were concentrated in the northwestern U.S. and many smaller fires were visible in other parts of the country (Fig. 10a). In the 1994–2003 time period (Fig. 10b), there was an increase in the overall amount of area burned, especially in the West. This trend continued into the 2004–2013 time period (Fig. 10c), but there were also a number of large fires that occurred in the Great Plains. Across all three time periods, eastern Kansas and the southeastern U.S. consistently had a large amount of area burned (Fig. 11).

Examining the trends in burned area in 10-year time intervals using the BAECV products, we found that area burned increased by 65% from 190,454 km<sup>2</sup> in 1984–1993 to 313,955 km<sup>2</sup> in 1994–2003, and then increased again by 37% to 431,649 km<sup>2</sup> in 2004–2013 (Table 9; Fig. 11). Between 1984–1993 and 1994–2003, rates of increase in burned area were greatest in the West (84%) and the Great Plains (70%) followed by the East (33%). However, between 1994–2003 and 2004–2013 the rates of change were greatest in the East (46%), followed by the Great Plains (44%) and then the West (27%).

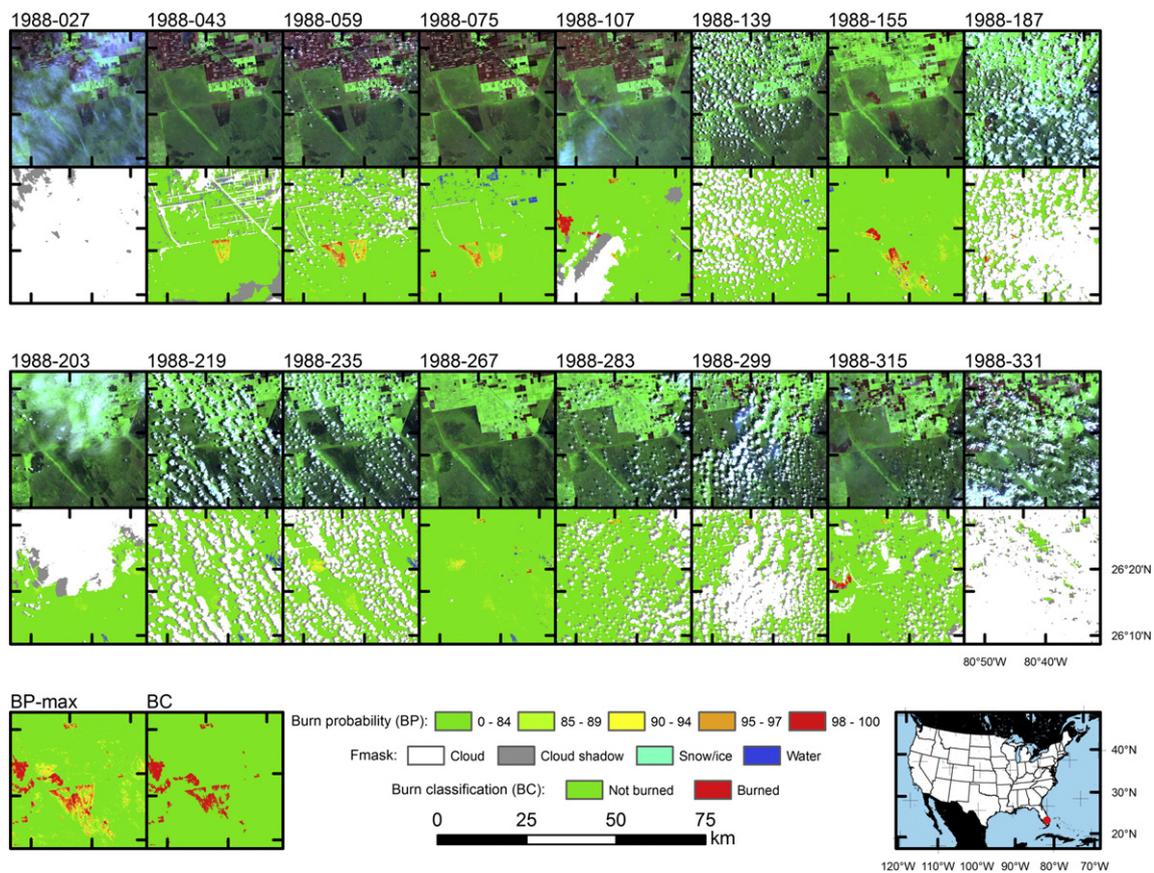
For many regions, rates of change in burned area estimated from the MTBS data were higher than they were from the BAECV data (Table 9; Fig. 11). Across CONUS, MTBS burned area increased by 73% between 1984–1993 and 1994–2003, and again by 66% between 1994–2003 and 2004–2013. Rates of change in burned area between 1984–1993 and 1994–2003 and between 1994–2003 and 2004–2013, respectively

were 95% and 24% in the West, 96% and 259% in the Great Plains, and –11% and 141% in the East.

### 4. Discussion

We designed and implemented the Landsat BAECV algorithm to extract burned areas from temporally-dense Landsat data across a diverse range of vegetation types. After implementation, we applied the BAECV algorithm to all available Landsat data covering the CONUS. To the best of our knowledge, this is the first time the entire archive of Landsat TM and ETM+ data has been processed for burned area mapping over such an extensive area. The BAECV products are currently the only moderate resolution data that characterize burned areas consistently over a long time period and large spatial extent in the CONUS. The products produced by the BAECV algorithm are available online at <http://dx.doi.org/10.5066/F73B5X76>.

Few algorithms exist that were specifically developed to automatically detect burned areas using large volumes of Landsat data. The Goodwin and Collett (2014) and the Boschetti et al. (2015) studies both identified areas of change and then used a series of rules to classify which regions of change were because of fires. The Boschetti et al. (2015) algorithm also incorporated MODIS active fire detections to separate burned areas from other types of change. Our approach is similar to the approach taken by Goodwin and Collett in that we both used lagged-summaries of Landsat data as a reference to measure change against in addition to a region-growing algorithm; however, our approach used longer lags and also incorporated a wider range of spectral indices as predictors. Like the Goodwin and Collett study, we also found



**Fig. 5.** Example of Landsat time series data and corresponding burn probability images for southern Florida used to generate the annual composite of maximum burn probability (BP-max) and the burn classification (BC). Images are labeled by year and Julian date.

that the thermal band was important for identifying burned areas and had difficulties consistently detecting burned areas in agricultural lands. The algorithm developed by [Boschetti et al. \(2015\)](#) was applied to 1 year of Landsat 7 ETM+ data for a subset of WRS2 path/rows over the western U.S.; areas that are primarily forested. Their approach is different in that it used a series of rules to classify pixels into spectral categories, identified pixels that changed spectral categories between weekly Landsat composites, and applied an additional set of rules to determine which changed pixels were potentially burned. Potentially burned pixels were then grouped and accepted or rejected as burned if they were in close proximity to MODIS active fires. Their approach benefits from using the MODIS active fire detections as an ancillary data source; however, this limits the applicability of their approach to time periods when MODIS or other active fire detections are available (e.g. 2000–present).

Similar to other algorithms ([Goodwin and Collett, 2014](#); [Boschetti et al., 2015](#)), the BAECV algorithm also incorporates a number of spectral indices to identify burned and unburned areas and is applicable to both forest and non-forest ecosystems. The motivation behind this

was based on the belief that a combination of spectral indices would more accurately allow us to identify burned areas than any single spectral index individually. Using the ecoregion, sensor type (TM or ETM+), spectral indices, their lagged summaries, and change metrics as predictor variables in a GBRM allowed for use of a large number of predictors without having to establish complicated rules for combining the results of separate classifiers using individual predictor variables.

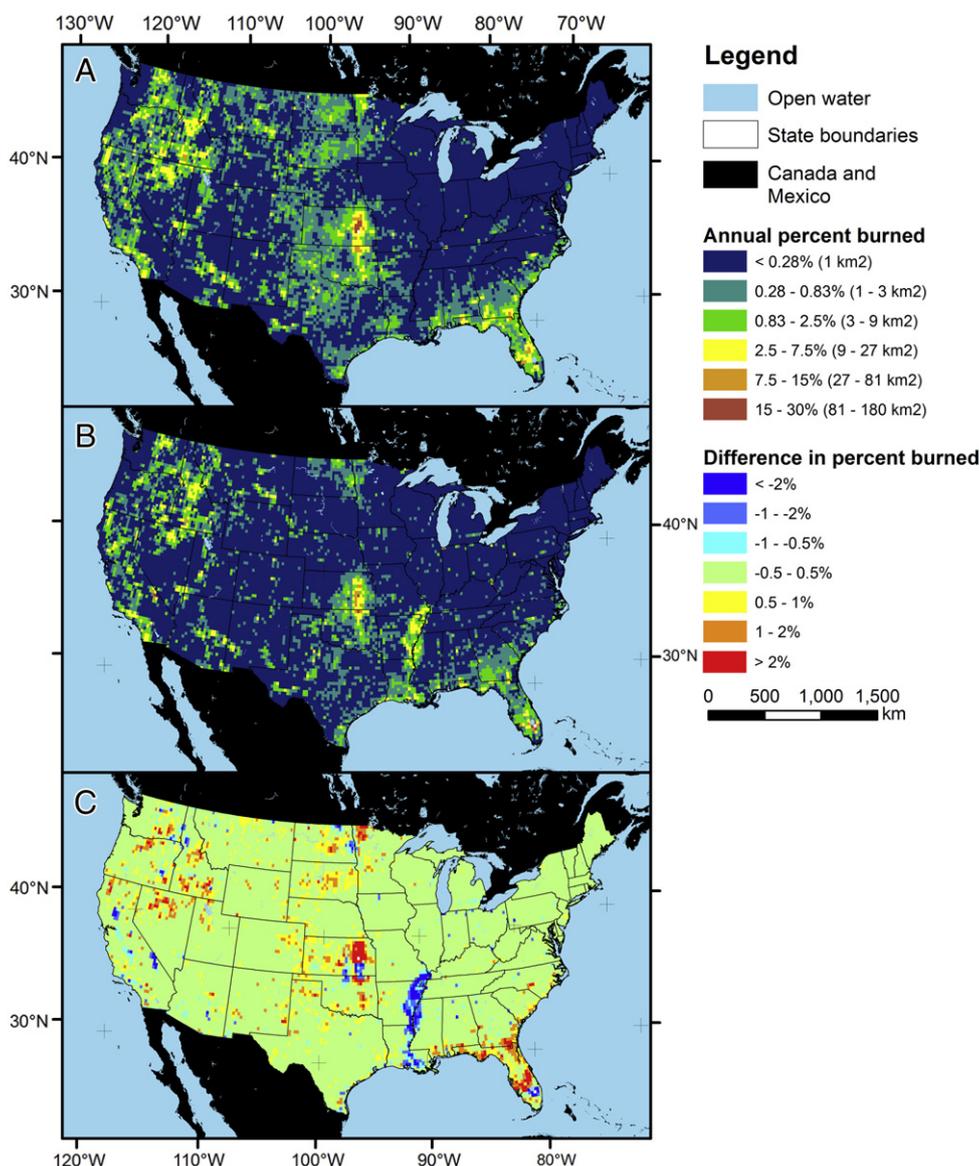
The relative importance values of the predictor variables demonstrated that characterizing land surface conditions in Landsat data both before and after fires is critical for mapping burned areas. However, we were surprised that the change metrics were not more important than they were, but perhaps this information is already well captured by the scene-level and reference predictors. We also examined the partial dependence plots of a few example predictors. These plots showed that Landsat bands 5 and 7, and the NBR2 index captured the spectral responses of burned areas that have been well documented in past studies ([Kasischke and French, 1995](#); [Rogan and Yool, 2001](#); [Trigg and Flasse, 2000](#)). The partial dependence plots of the 3-year lagged mean and standard deviation of NBR2 demonstrated that burned pixels were more likely where the 3-year lagged mean of NBR2 was above 0.2 and less likely as the 3-year lagged standard deviation of NBR2 increased. These responses suggest that the lagged predictor variables were keying in on pre-fire land surface conditions and filtering wildland vegetation capable of burning from other land cover types.

Compared to existing data and studies, one advantage of the BAECV products is that they provide the longest consistent record of burned area information for the CONUS, 32 years. That record will grow as the BAECV algorithm is updated to include data from OLI and potentially other sensors. Furthermore, if automated processes to atmospherically correct and generate cloud masks were operational for MSS data, our approach could also be expanded to include another 10 years of fire

**Table 5**

Total area burned from the Landsat Burned Area Essential Climate Variable (BAECV) and the Global Fire Emissions Database (GFED) from 1997 to 2015, and percent difference between the BAECV and GFED data.

Region	Burned area (km <sup>2</sup> )		Percent difference
	BAECV	GFED	
West	286,930	217,652	32%
Great Plains	165,240	171,620	−4%
East	270,717	143,720	88%
CONUS	722,887	532,992	36%



**Fig. 6.** One-quarter degree summaries of mean annual burned area from 1984 to 2013 for the (A) Burned Area Essential Climate Variable (BAECV) burn classification product, (B) the Global Fire Emissions Database (GFED), and (C) their difference (BAECV - GFED).

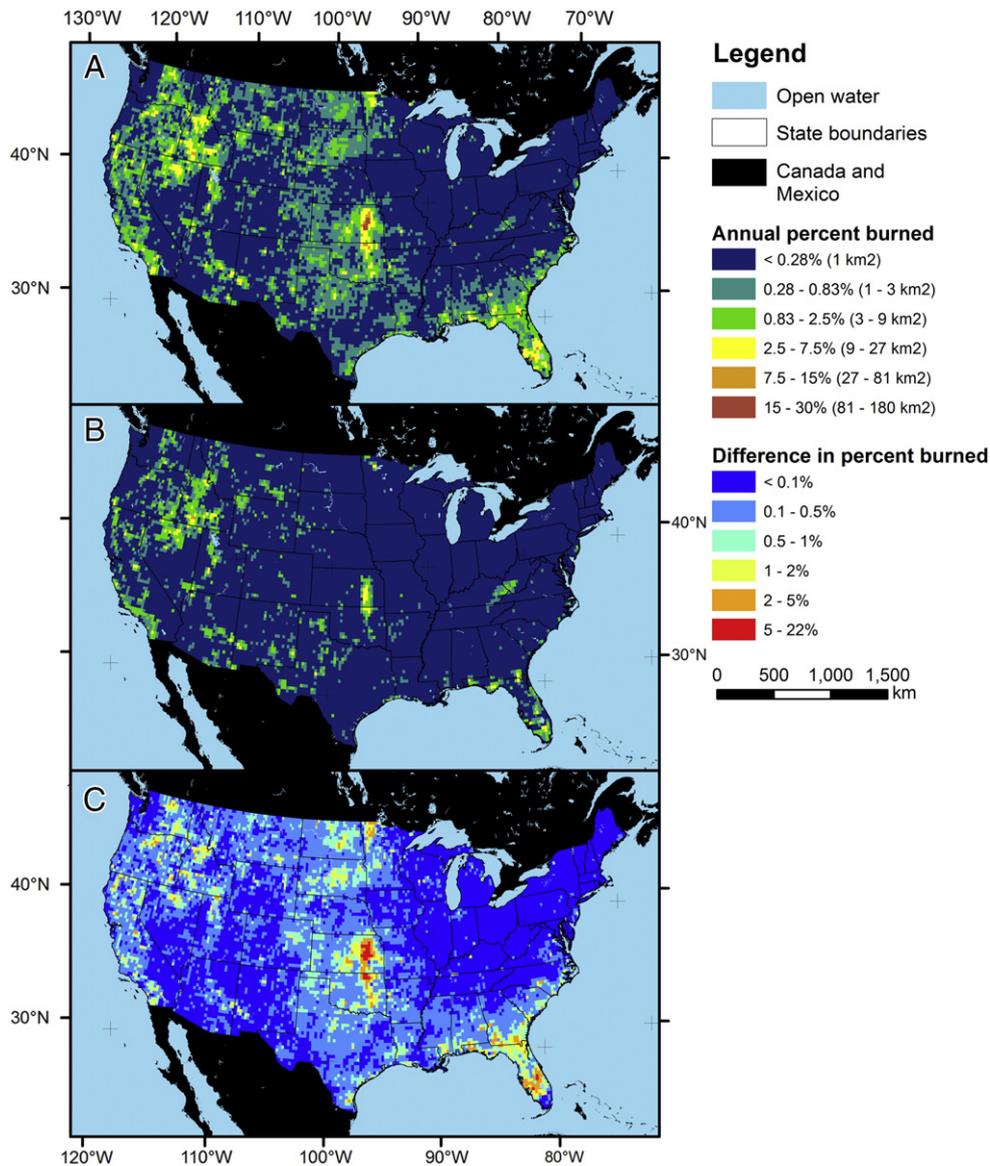
history (1972–1983). However, the predictor variables we used would be limited to those that can be derived from the MSS spectral bands. This has promise as recent studies have improved georeferencing (Devaraj and Shah, 2014; Kennedy and Cohen, 2003) and automating cloud masking of MSS data (Braaten et al., 2015), and radiometric calibration (Markham and Helder, 2012); however, processes like these and a standardized method to atmospherically correct MSS data have not yet been implemented U.S. Geological Survey's ESPA Landsat data ordering interface.

#### 4.1. Differences between BAECV, GFED, and MTBS data

Comparison with existing datasets is an important process to help potential users understand what is novel and unique about the BAECV data. We did this using the GFED, used globally to monitor patterns and trends in burned area and emissions, and also the MTBS data, which are well known by scientists and land managers in the United States. We found that the burned area mapped by BAECV products was more similar to the GFED products (1997–2015) than the MTBS data (1984–2013). However, the spatial patterns of burned area were different between the BAECV and GFED products. Most notably, GFED

appears to be including more burned area in the Central California Valley, the southern Flint Hills, the Mississippi Alluvial Plain, and the Southern Florida Coastal Plain ecoregions (Fig. 6). Other than the Flint Hills, these ecoregions are dominated by agriculture and agricultural fires are common. The BAECV likely missed these burned areas because it is challenging to identify burned areas with the relatively long revisit intervals of the Landsat sensors and difficulties distinguishing burned areas from agricultural tillage. Ecoregions where the BAECV products mapped more burned area than the GFED data included the Columbia Plateau, the Central and Northern Basin and Range, the northern Flint Hills and other parts of the Great Plains, and the Southern Coastal Plain ecoregions (Fig. 6). Some of these differences might be caused by BAECV commission errors, especially in agricultural lands. However, these are also ecoregions where fires are common, spread quickly, and leave little residual heat. Because of this, they might be missed by active fire detection algorithms developed for sensors like MODIS (Hawbaker et al., 2008). Consequently, we might expect the GFED to underestimate burned area in these ecoregions and also explain some of the differences we observed.

In contrast to the differences between the BAECV products and GFED, the BAECV products documented spatial and temporal patterns



**Fig. 7.** One-quarter degree summaries of mean annual burned area from 1984 to 2013 for the (A) Burned Area Essential Climate Variable (BAECV) burn classification product, (B) Monitoring Trends in Burn Severity (MTBS) data, and (C) their difference (BAECV – MTBS).

of burned area that are regionally different from existing long-term burned area information included in the MTBS data and other fire occurrence datasets that have inconsistent reporting effort over time (Brown et al., 2002; Short, 2014; Short, 2015). In general, the BAECV and MTBS data had similar amounts of burned area in the West. However, the BAECV products included a greater amount of burned area than MTBS in the Great Plains and the East, both areas where our understanding of long-term fire patterns is incomplete.

**Table 6**

Total area burned from the Landsat Burned Area Essential Climate Variable (BAECV) and the Monitoring Trends in Burn Severity Project (MTBS) from 1984 to 2013, and percent difference between the BAECV and MTBS data.

Region	Burned area (km <sup>2</sup> )		Percent difference
	BAECV	MTBS	
West	369,157	282,611	31%
Great Plains	342,372	83,092	312%
East	224,528	67,432	233%
CONUS	936,057	433,135	116%

Burned areas not included in the MTBS dataset primarily accounted for the large differences between the MTBS and BAECV products. The additional burned areas mapped by the BAECV occurred in a range of land cover types, including forest (Fig. 12a), shrublands (Fig. 12a, d), and grasslands (Fig. 12c, d), wetlands (Fig. 12b and d), and agriculture (Fig. 12b). Many of the BAECV burned areas had no corresponding MTBS perimeter but were near MODIS active fire detections, adding confidence to our results. However, many other BAECV burned areas without MTBS perimeters were also distant from MODIS active fire detections. Some of these burned areas had patch shapes suggesting they were fires. Others had more rectilinear shapes following the boundaries of agricultural fields and/or pastures. Burning in these land cover types is not uncommon (McCarty et al., 2009), but difficult to determine with confidence with Landsat data. Similarly, BAECV burned areas were occasionally mapped in clear-cut forests, especially in the southeastern US. It is challenging to determine whether or not these were errors as burning is sometimes used as a management action after harvest to remove litter and duff, reduce understory competition, promote seed germination, and growth (Ryan et al., 2013). In other cases, visual comparison of the BAECV products with the original Landsat imagery, MODIS active fires, and federal fire occurrence databases occasionally showed

**Table 7**

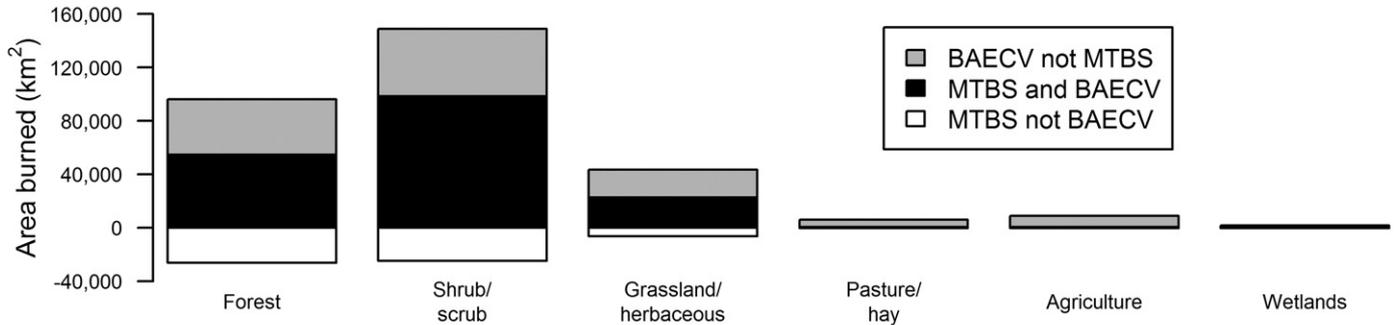
Total area and burned area for the conterminous United States (CONUS) and regions of CONUS from 1984 through 2013, as well as, comparisons in the percentages of burned area found between the Landsat Burned Area Essential Climate Variable (BAECV) and the Monitoring Trends in Burn Severity (MTBS) products.

Region	BAECV burned area (km <sup>2</sup> )	MTBS burned area (km <sup>2</sup> )	Mapped by BAECV but not MTBS (km <sup>2</sup> )	Mapped by BAECV and MTBS (km <sup>2</sup> )	Mapped by MTBS but not BAECV (km <sup>2</sup> )	Percent of BAECV not in MTBS	Percent of MTBS found by BAECV
West	369,157	282,611	164,228	204,929	77,682	44%	73%
Great Plains	342,372	83,092	285,833	57,539	25,552	83%	69%
East	224,528	67,432	195,244	29,284	38,148	87%	43%
CONUS	936,057	433,135	645,305	291,752	141,382	69%	67%

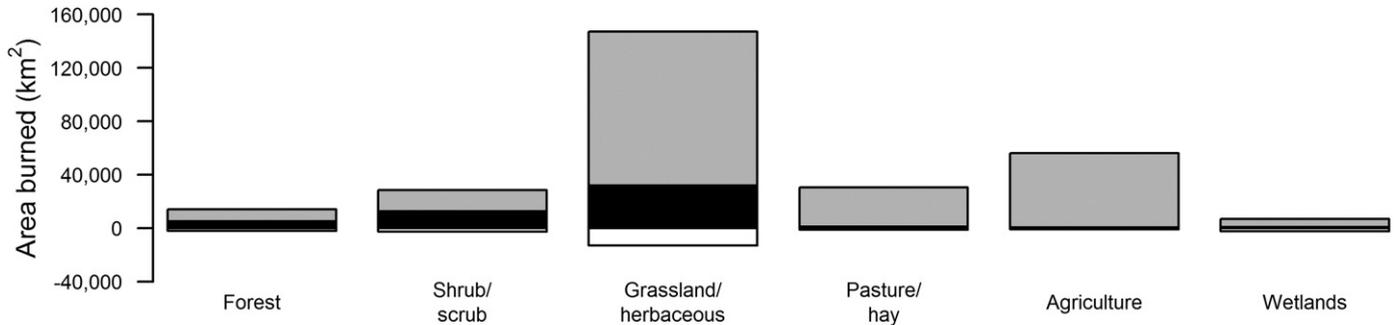
burned areas that were misclassified by the BAECV algorithm, especially in areas with drought-induced vegetation senescence or in scenes with georegistration errors higher than those reported in their metadata. In these cases, false changes could have triggered our algorithm. Separating incorrectly identified burned areas from correctly identified areas in the BAECV data remains challenging in these areas.

Even though the BAECV products identified much more burned area across the CONUS than MTBS did, the BAECV products did not map all burned areas included in the MTBS data. This was especially true in the East, where the BAECV products missed 57% of the MTBS burned area. The differences were less pronounced in the Great Plains and the West, where the BAECV products missed 31% and 27% of the MTBS burned area, respectively. These omission errors could have been

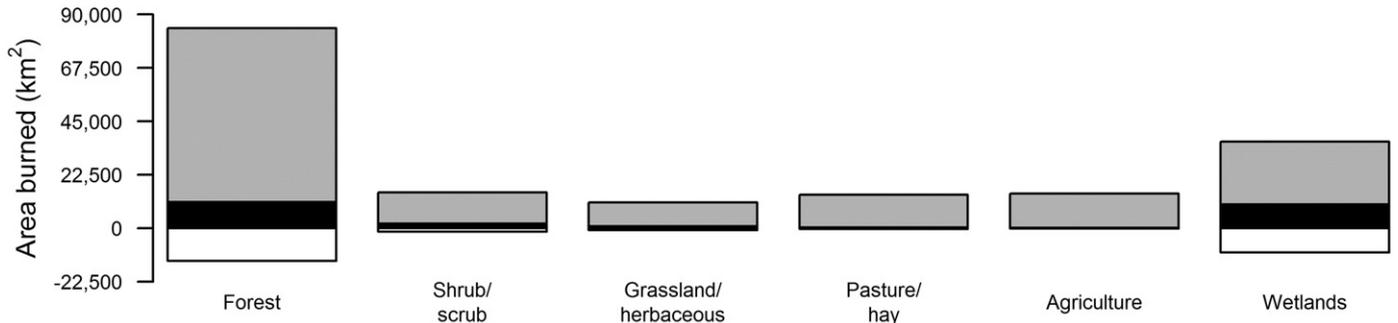
**A) West**



**B) Great Plains**



**C) East**



**Fig. 8.** The amount of area burned from 1992 through 2013 common and unique to the Landsat Burned Area Essential Climate Variable (BAECV) and the Monitoring Trends in Burn Severity (MTBS) data by regions of the conterminous United States. X-axis labels are generalized classes from the National Land Cover Database.

**Table 8**

The percent of Monitoring Trends in Burn Severity (MTBS) fire severity pixels detected by the Burned Area Essential Climate Variable (BAECV). All MTBS low, moderate and high severity burn pixels were compared to the annual BAECV product over the conterminous U.S.

Region	Severity (%)		
	Low	Moderate	High
East	46%	71%	87%
Great Plains	74%	90%	89%
West	78%	88%	85%
CONUS	70%	87%	85%

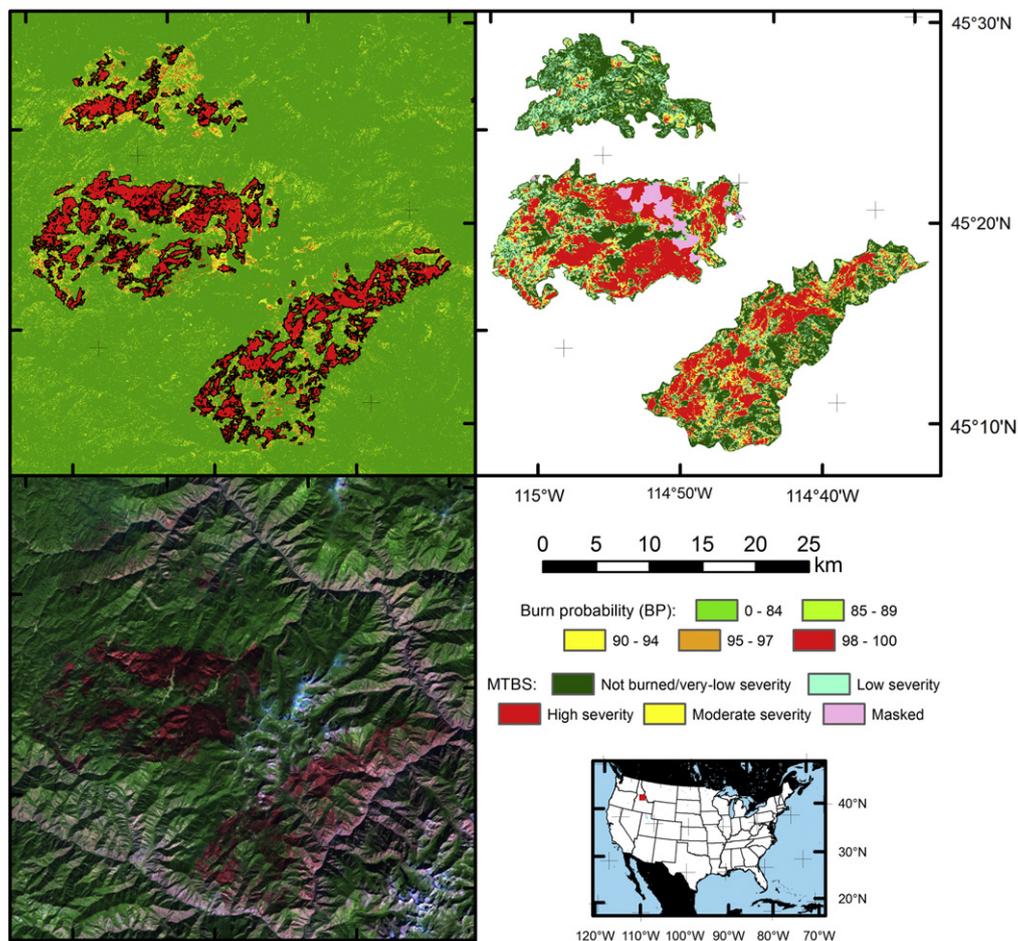
caused for a variety of reasons. Cloud cover could have potentially obscured burned areas in regions with rapid post-fire vegetation recovery (e.g. the Great Plains and the East). Conservative cloud masks generated by FMask could have occluded these fires from the BAECV algorithm, even if the burned areas were visible in the Landsat images (for example see the 1988-027 Landsat image in Fig. 5). Additional omission errors could have been caused by challenges with identifying low-severity burns, and understory burns that left forest canopies intact. In these areas, land-surface changes may not have been intense enough for detection by the BAECV algorithm and may have caused confusion for the MTBS analysts – MTBS may overestimate burned area as commission error for burned area has been found to be up to 24% (Meddens et al., 2016). This was likely the case in the East, where we found that the BAECV products had 46% omission error rates for low-severity burns. Therefore, our classification results may be improved by refining our training data and visually verifying

that pixels identified as burned by the MTBS data were actually visible as burned in the Landsat images.

Another reason for the differences between the BAECV and MTBS data may be related to the MTBS methodology. It prioritizes large fires ( $\geq 2$  km<sup>2</sup> in the eastern U.S. and  $\geq 4$  km<sup>2</sup> in the western U.S.) reported in federal fire occurrence databases. The reported point locations and dates for those fires are then used to identify pre- and post-fire Landsat scenes to use for perimeter and burn severity mapping (Eidenshink et al., 2007). Thus, the MTBS data are most incomplete earliest in the dataset when federal fire occurrence reporting was most inconsistent (Brown et al., 2002) and in areas where fires are not reported in the federal fire occurrence databases. This is especially likely for the Great Plains and the East where there are many fires but little federally owned land. In addition to reliance on federal fire occurrence databases, some fires may be missing from the MTBS data because of the history of the project; MTBS stopped mapping prescribed fires on non-federal lands in 2014 because it was not feasible to map the large number of reported prescribed fires in the Great Plains and the East. Additionally, the MTBS project started in 2006, before the Landsat archive was open and free, thus fewer scenes could be purchased and some fires could not be mapped. Consequently, estimates of rates of change in burned area will be amplified because of underestimates of burned area in the early years (pre-2004) of the MTBS dataset.

#### 4.2. Validation of the BAECV products

Completing a thorough validation of the BAECV products was beyond the scope of this study and not practical because the MTBS data



**Fig. 9.** A comparison between the Burned Area Essential Climate Variable (BAECV) (top left), the Monitoring Trends in Burn Severity (MTBS) mapped fire severity (top right), and the Landsat reference dataset (bottom left) for evergreen forest in central Idaho (path 41, rows 28 and 29, Julian day 273, 1988). This example is typical in that the BAECV burn area extent is often more conservative relative to the MTBS data.

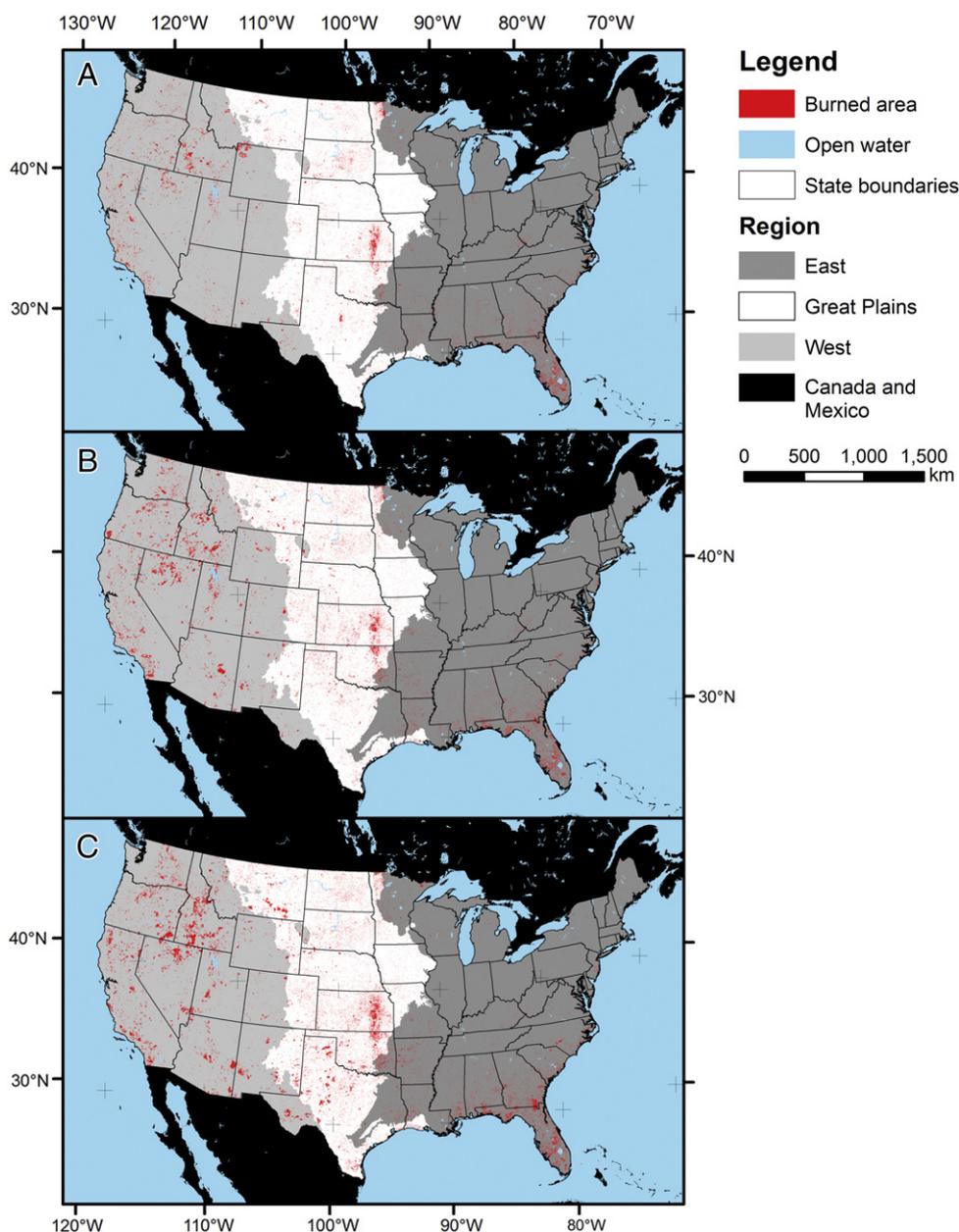


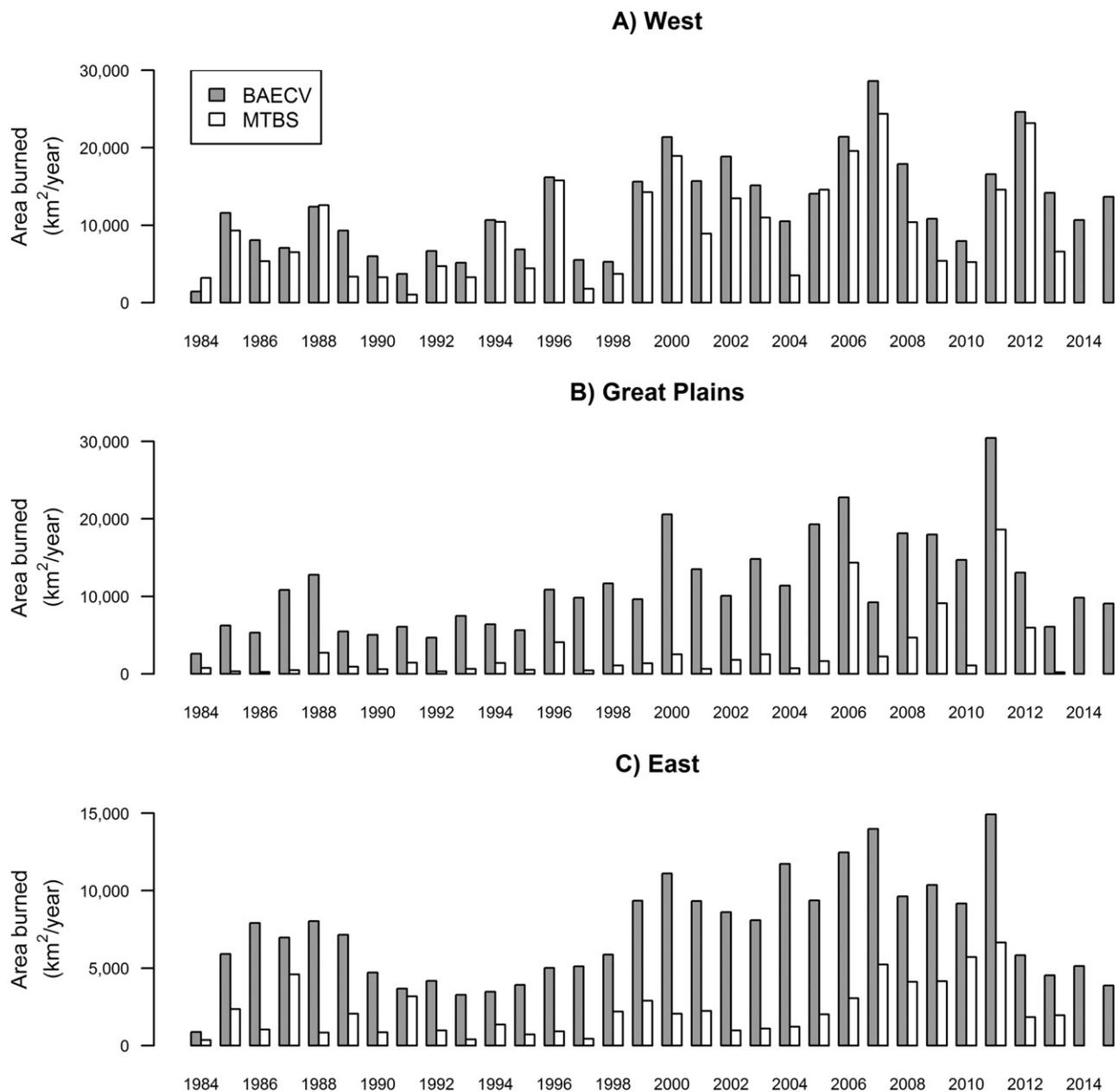
Fig. 10. Burned areas for (A) 1984–1993, (B) 1994–2003, and (C) 2004–2013 from the Landsat Burned Area Essential Climate Variable.

are known to be incomplete. However, validation is essential to help potential users assess whether or not products will meet their requirements for use and how to interpret findings based on them (Morissette et al., 2006). Instead, we devoted an entire companion paper to validation of the BAECV products (Vanderhoof et al., 2017). The methods used in this validation effort largely follow the guidelines established by the Committee on Earth Observation Satellites (CEOS), Land Product Validation Subgroup (LPVS) and followed by the fire\_cci project under the European Space Agency's (ESA) Climate Change Initiative (CCI) (Padilla et al., 2015). However, the approach taken by Vanderhoof et al. is unique in that 3 image analysts each independently derived validation data from Landsat imagery in 28 path/rows for five years (1988, 1993, 1998, 2003, and 2008). They analyzed their results using regions similar to those used in this study (East, Great Plains, and West), but subdivided the West into the Arid West and Mountain West. They found that omission and commission error rates were balanced when reference data from 2 analysts were used. Overall accuracy for burned areas in the BAECV products was found to be very high; 99.9% for CONUS, the East, Great Plains,

and Arid West, and dropped to 99.8% for the Mountain West. Omission and commission error rates were more variable. Omission rates ranged from 31% in the Arid West, 41% in the Mountain West, 62% in the Great Plains, 67% in the East, and 43% across CONUS. Commission rates ranged from 24% in the Arid West, 32% in the Mountain West, 57% in the Great Plains, 47% in the East, and 34% across CONUS. These error rates are larger than those in previous studies using individual scenes or with limited spatial extent (e.g., Bastarrika et al., 2011; Chuvieco et al., 2002; Goodwin and Collett, 2014; Koutsias, 2003; Stroppiana et al., 2012), but lower than those documented for global coarse-resolution products (Padilla et al., 2015). The BAECV error rates are similar to those produced by VCT and LandTrendr, which are constrained to forested cover types (Thomas et al., 2011; Kennedy et al., 2015).

#### 4.3. Potential applications and use of the BAECV products

There is a wide range of potential applications for the BAECV products. Although an independent accuracy assessment of the BAECV



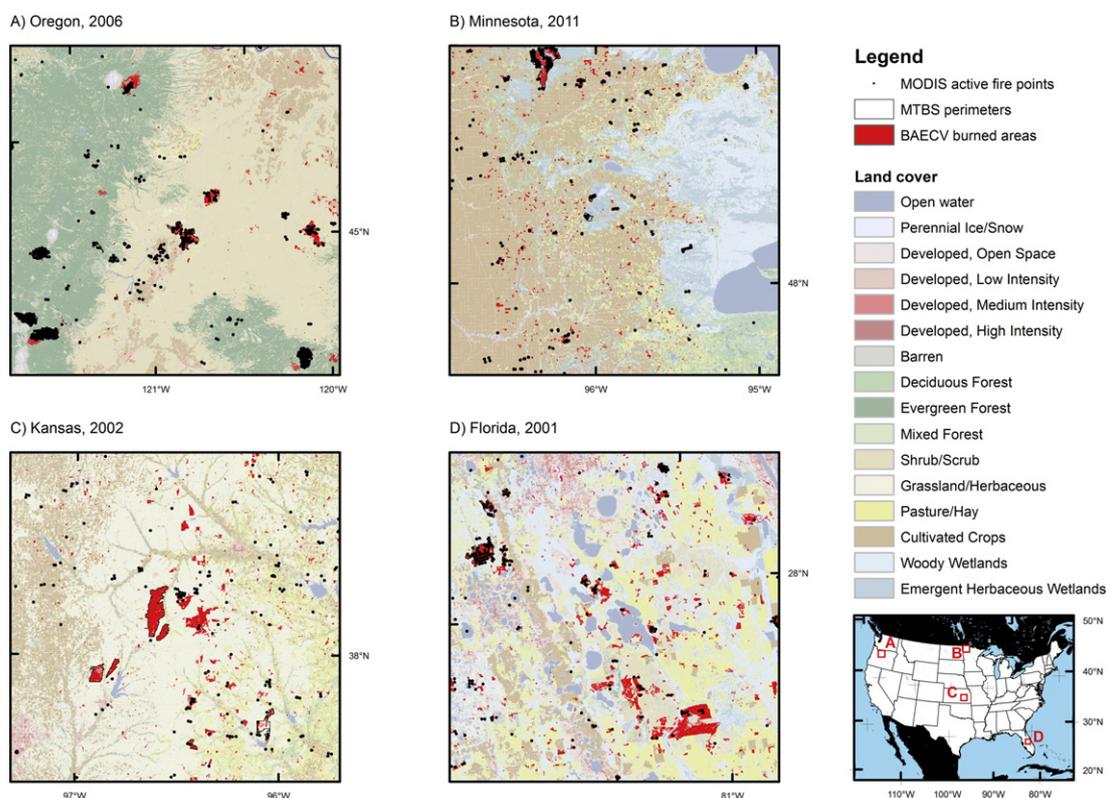
**Fig. 11.** Area burned by region and year as measured by the Landsat Burned Area Essential Climate Variable (BAECV) and the Monitoring Trends in Burn Severity Project (MTBS) from 1984 through 2015. MTBS data were not available for 2014 and 2015.

products has been completed (Vanderhoof et al., 2017), users should assess the data to determine if they are appropriate and have acceptable accuracy levels for their needs. For analyses covering large spatial and temporal extents, we recommend using the data as they are, as errors of omission and commission for burned areas are approximately balanced across the CONUS (Vanderhoof et al., 2017), but potentially

masking burned areas in agricultural lands, since this land cover type has been shown to experience the highest rates of error, relative to other land cover types (Vanderhoof et al., 2017). The BAECV certainly detects burned areas in agricultural lands, but not consistently because of Landsat's 16-day revisit interval and the large amount of variability in vegetation and soil condition in agricultural lands. For smaller-scale

**Table 9**  
Area burned by region and decade-long time periods as measured by the Landsat Burned Area Essential Climate Variable (BAECV) and the Monitoring Trends in Burn Severity Project (MTBS) from 1984 through 2013. Percentages indicate change relative to the previous time period.

Region	ECV burned area (km <sup>2</sup> )			MTBS burned area (km <sup>2</sup> )						
	1984–1993	1994–2003	2004–2013	1984–1993	1994–2003	2004–2013				
West	71,354	131,168	84%	166,635	27%	52,536	102,669	95%	127,406.5	24%
Great Plains	66,419	112,933	70%	163,020	44%	8305	16,292	96%	58,493.8	259%
East	52,680	69,855	33%	101,993	46%	16,643	14,872	–11%	35,916.2	141%
CONUS	190,454	313,955	65%	431,649	37%	77,484	133,833	73%	221,817	66%



**Fig. 12.** Examples of burned areas mapped by the Burned Area Essential Climate Variable (BAECV) algorithm, Monitoring Trends in Burn Severity (MTBS) project, and the Moderate Resolution Imaging Spectroradiometer (MODIS) for select areas in the conterminous United States.

analyses, we would advise users to visually inspect the burned areas mapped by the BAECV algorithm, compare them to Landsat or other imagery, and manually edit as needed.

In addition to the long time series provided by the BAECV products, their moderate spatial resolution (30 m) also makes them amenable for ecological studies. Few other burned area products exist at this resolution, except for the MTBS data which are known to be incomplete. Coarse-resolution burned area data derived from MODIS or VIIRS imagery do not offer the detail required for on-the-ground ecological applications (Kennedy et al., 2014). Having finer resolution data on burned areas will also facilitate spatial analysis assessing how fire occurrence varies among land cover and vegetation types and quantifying fire impacts like emissions modeling. These types of analysis have been limited because coarse-resolution burned area data lack small fires (Randerson et al., 2012) and make it difficult to assess exactly what was burning in highly heterogeneous landscapes (Eva and Lambin, 2000). The moderate spatial resolution of the BAECV products should facilitate investigations relating patterns of burning to other moderate-resolution data, for instance land cover or biomass consumption.

We designed our algorithm to be flexible enough to be applied to other parts of the world beyond the U.S. However, the availability of Landsat images will limit the temporal span of our results. In many places of the world, image availability is best after Landsat ETM+ became operational in 1999 and collection of both TM and ETM+ images was more systematic (Kovalskyy and Roy, 2013). For time periods before 1999, the reference and change metrics we derive from historic Landsat imagery will need to be reassessed and the lag times over which they are calculated may need to be increased to ensure enough images are incorporated to capture the range of variability in pre-fire surface conditions. Given data availability limitations, the BAECV algorithm could potentially be adjusted to rely more on scene-level predictors than on reference and change predictors. The importance of the scene-level band values for our algorithm (Fig. 2) indicates that this may be possible, but we would anticipate higher commission rates

without variables representing pre-fire conditions and change from pre-fire conditions. Resulting BAECV products prior to 1999 might require review prior to use in analyses to remove or reduce commission errors. An additional limitation to overcome would be the availability of training and validation data. Few areas outside of the U.S. have data that represent heterogeneity within fire perimeters with the level of detail that the MTBS data provide. Because the BAECV algorithm requires a large sample of burned and unburned locations to work effectively, new data would need to be collected before training and applying the BAECV algorithm elsewhere. However, that constraint limits wide-spread application of almost all change-detection and burned area mapping algorithms as well.

## 5. Conclusion

We developed and implemented the Landsat BAECV algorithm to identify burned areas in temporally-dense time series of Landsat images for the CONUS and to contribute USGS's efforts to produce application-ready Landsat science products and to GCOS efforts to produce global burned area ECVs. The data products produced by the BAECV algorithm document patterns of fire occurrence that are not well characterized by existing fire datasets in the U.S. and our approach can be extended to other regions of the world with some extra effort. These data could help to better understand past patterns of fire occurrence, the drivers that created them, and the impacts fires have on natural and human systems.

## Data and resources

BAECV outputs including annual mosaics for the CONUS of maximum burn probability and burn classification are available at <http://dx.doi.org/10.5066/F73B5X76>. We encourage interested users to download and evaluate the BAECV outputs and to contact us at [BAECV@usgs.gov](mailto:BAECV@usgs.gov) with suggestions and feedback.

## Acknowledgements

We are grateful to the U.S. Geological Survey (Project number RL00F460100), Climate and Land Use Mission Area, Land Remote Sensing Program for providing funding to support this research. This work benefitted from the high-quality burned area data produced from the Monitoring Trends in Burn Severity (MTBS) Project. Hayley Distler and Matthew Galbraith helped to generate and visually assess the Landsat Burned Area Essential Climate Variable data for the conterminous United States. We are also grateful to Carol Mladinich and the MTBS and LANDFIRE teams at the U.S. Geological Survey Earth Resources Observation Science Center for assessing initial results of the algorithm and providing feedback to help improve initial versions of the algorithm. The editor, managing editor, and anonymous reviewers for Remote Sensing of Environment, in addition to Janet Slate, Eugene Ellis, Randall Schumann, Eugene Schweig, Kurtis Nelson, Elizabeth Lile, and Marilyn Gibbons who all provided insightful comments on a previous draft of this manuscript and their comments helped greatly to improve the completeness and clarity of the manuscript.

The BAECV algorithm was implemented in Python v. 2.7.5 (<http://www.python.org>; Python Software Foundation, 2016), with the NumPy 1.9.2 (<http://www.numpy.org>; van der Walt et al., 2011), SciPy 0.16.0 (<http://www.scipy.org>; Millman and Aivazis, 2011; Oliphant, 2007), Scikit-learn 0.16.1 (<http://scikit-learn.org>; Pedregosa et al., 2011), Scikit-image v. 0.11.3 (<http://scikit-image.org>; van der Walt et al., 2014), and the Geospatial Data Abstraction Library v. 1.11.2 (<http://www.gdal.org>; Open Source Geospatial Foundation, 2016) packages. Analysis of BAECV products for this study were completed using the R language and environment for statistical computing (<https://www.r-project.org>; R Foundation for Statistical Computing, 2014), v 3.1.1 and ArcGIS Desktop v. 10.3.1 (<http://www.esri.com/>; Environmental Systems Research Institute, 2015). Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

## References

- Abatzoglou, J.T., Williams, A.P., 2016. Impact of anthropogenic climate change on wildfire across western US forests. *Proc. Natl. Acad. Sci.* 113, 11770–11775.
- Alonso-Canas, L., Chuvieco, E., 2015. Global burned area mapping from ENVISAT-MERIS and MODIS active fire data. *Remote Sens. Environ.* 163, 140–152.
- Bachelet, D., Neilson, R.P., Hickler, T., Drapek, R.J., Lenihan, J.M., Sykes, M.T., Smith, B., Sitch, S., Thonicke, K., 2003. Simulating past and future dynamics of natural ecosystems in the United States. *Global Biogeochem. Cycles* 17:1045. <http://dx.doi.org/10.1029/2001GB001508>.
- Bastarrica, A., Chuvieco, E., Martin, M.P., 2011. Mapping burned areas from Landsat TM/ETM plus data with a two-phase algorithm: balancing omission and commission errors. *Remote Sens. Environ.* 115, 1003–1012.
- Boschetti, L., Roy, D.P., Justice, C.O., Humber, M.L., 2015. MODIS-Landsat fusion for large area 30 m burned area mapping. *Remote Sens. Environ.* 161, 27–42.
- Braaten, J.D., Cohen, W.B., Wang, Z., 2015. Automated cloud and cloud shadow identification in Landsat MSS imagery for temperate ecosystems. *Remote Sens. Environ.* 169:128–138. <http://www.sciencedirect.com/science/article/pii/S0034425715300948>.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. *Classification and Regression Trees*. 1st ed. Chapman and Hall, New York, NY, USA.
- Brooks, E.B., Wynne, R.H., Thomas, V.A., Blinn, C.E., Coulston, J.W., 2014. On-the-fly massively multitemporal change detection using statistical quality control charts and Landsat data. *IEEE Trans. Geosci. Remote Sens.* 52, 3316–3332.
- Brown, T.J., Hall, B.L., Mohrle, C.R., Reinbold, H.J., 2002. *Coarse Assessment of Federal Wildland Fire Occurrence Data: Report for the National Wildfire Coordinating Group*. Desert Research Institute, Reno, NV, USA.
- Chan, J.C.W., Chan, K.P., Yeh, A.G.O., 2001. Detecting the nature of change in an urban environment: a comparison of machine learning algorithms. *Photogramm. Eng. Remote Sens.* 67, 213–225.
- Chuvieco, E., Congalton, R.G., 1988. Mapping and inventory of forest fires from digital processing of TM data. *Geocarto Int.* 4, 41–53.
- Chuvieco, E., Martín, M.P., Palacios, A., 2002. Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination. *Int. J. Remote Sens.* 23, 5103–5110.
- Coulston, J.W., Moisen, G.G., Wilson, B.T., Finco, M.V., Cohen, W.B., Brewer, C.K., 2012. Modeling percent tree canopy cover: a pilot study. *Photogramm. Eng. Remote Sens.* 78, 715–727.
- Crist, E.P., 1985. A TM tasseled cap equivalent transformation for reflectance factor data. *Remote Sens. Environ.* 17, 301–306.
- Dennison, P.E., Brewer, S.C., Arnold, J.D., Moritz, M.A., 2014. Large wildfire trends in the western United States, 1984–2011. *Geophys. Res. Lett.* 41, 2928–2933.
- Devaraj, C., Shah, C.A., 2014. Automated geometric correction of Landsat MSS L1G imagery. *IEEE Geosci. Remote Sens. Lett.* 11, 347–351.
- Dozier, J., 1981. A method for satellite identification of surface-temperature fields of subpixel resolution. *Remote Sens. Environ.* 11, 221–229.
- Eidenshink, J., Schwind, B., Brewer, K., Zhu, Z.L., Quayle, B., Howard, S., 2007. A project for monitoring trends in burn severity. *Fire Ecol.* 3, 3–21.
- Elvidge, C.D., Kroehl, H.W., Kihn, E.A., Baugh, K.E., Davis, E.R., Hao, W.M., 1996. Algorithm for the retrieval of fire pixels from DMSP Operational Linescan System. In: Levine, J.S. (Ed.), *Global Biomass Burning*. MIT Press, Cambridge, MA, USA, pp. 77–85.
- Environmental Systems Research Institute, 2015. *ArcGIS Desktop: Release 10.3.1* (Redlands, CA).
- Eva, H., Lambin, E.F., 1998. Burnt area mapping in Central Africa using ATSR data. *Int. J. Remote Sens.* 19, 3473–3497.
- Eva, H., Lambin, E.F., 2000. Fires and land-cover change in the tropics: a remote sensing analysis at the landscape scale. *J. Biogeogr.* 27, 765–776.
- Flannigan, M.D., Vonder Haar, T.H., 1986. Forest-fire monitoring using NOAA satellite AVHRR. *Can. J. For. Res.* 16, 975–982.
- Food and Agriculture Organization of the United Nations (2008). *Terrestrial Essential Climate Variables for Climate Change Assessment, Mitigation and Adaptation*; Global Terrestrial Observing System GTOS 52 - Biennial Report Supplement. In R. Sessa, & H. Dolman (Eds.). (Rome, Italy).
- Fraser, R.H., Li, Z., Cihlar, J., 2000. Hotspot and NDVI differencing synergy (HANDS): a new technique for burned area mapping over boreal forest. *Remote Sens. Environ.* 74, 362–376.
- French, N.H.F., McKenzie, D., Erickson, T., Koziol, B., Billmire, M., Endlsey, K.A., Yager Scheinerman, N.K., Jenkins, L., Miller, M.E., Ottmar, R.D., Prichard, S., 2014. Modeling regional-scale wildland fire emissions with the Wildland Fire Emissions Information System. *Earth Interact.* 18, 1–26.
- Friedl, M.A., Brodley, C.E., 1997. Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* 61, 399–409.
- Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N., Wickham, J., 2011. Completion of the 2006 National Land Cover Database for the conterminous United States. *Photogramm. Eng. Remote Sens.* 77, 858–864.
- Gao, B., 1996. NDWI - a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* 58, 257–266.
- Giglio, L., Kendall, J.D., Tucker, C.J., 2000. Remote sensing of fires with the TRMM VIRS. *Int. J. Remote Sens.* 21, 203–207.
- Giglio, L., Descloitres, J., Justice, C.O., Kaufman, Y.J., 2003. An enhanced contextual fire detection algorithm for MODIS. *Remote Sens. Environ.* 87, 273–282.
- Giglio, L., Loboda, T., Roy, D.P., Quayle, B., Justice, C.O., 2009. An active-fire based burned area mapping algorithm for the MODIS sensor. *Remote Sens. Environ.* 113, 408–420.
- Giglio, L., Randerson, J.T., van der Werf, G.R., 2013. Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). *J. Geophys. Res. Biogeosci.* 118, 317–328.
- Global Climate Observing System, 2004. *Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC, GCOS-92*. World Meteorological Organization, Geneva, Switzerland, p. 136.
- Goodwin, N.R., Collett, L.J., 2014. Development of an automated method for mapping fire history captured in Landsat TM and ETM plus time series across Queensland, Australia. *Remote Sens. Environ.* 148, 206–221.
- Hall, D.K., Ormsby, J.P., Johnson, L., Brown, J., 1980. Landsat digital analysis of the initial recovery of burned tundra at Kokolik River, Alaska. *Remote Sens. Environ.* 10, 263–272.
- Hansen, M.C., Egorov, A., Potapov, P.V., Stehman, S.V., Tyukavina, A., Turubanova, S.A., Roy, D.P., Goetz, S.J., Loveland, T.R., Ju, J., Kommareddy, A., Kovalsky, V., Forsyth, C., Bents, T., 2014. *Monitoring conterminous United States (CONUS) land cover change with Web-Enabled Landsat Data (WELD)*. *Remote Sens. Environ.* 140, 466–484.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. *The Elements of Statistical Learning; Data Mining, Inference, and Prediction*. 2nd ed. Springer, New York, NY, USA.
- Hawbaker, T.J., Zhu, Z., 2012. Projected future wildland fires and emissions for the Western United States, chapter 8 of Z. In: Zhu, Reed, B. (Eds.), *Baseline and Projected Future Carbon Storage and Greenhouse-Gas Fluxes in Ecosystems of the Western United States*: U.S. Geological Survey Professional Paper 1797 192 p. (Also available at <http://pubs.usgs.gov/pp/1797/>).
- Hawbaker, T.J., Radeloff, V.C., Syphard, A.D., Zhu, Z., Stewart, S.I., 2008. Detection rates of the MODIS active fire product in the United States. *Remote Sens. Environ.* 112, 2656–2664.
- Hawbaker, T.J., Radeloff, V.C., Stewart, S.I., Hammer, R.B., Keuler, N.S., Clayton, M.K., 2013. Human and biophysical influences on fire occurrence in the United States. *Ecol. Appl.* 23, 565–582.
- Hollmann, R., Merchant, C.J., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., Chuvieco, E., Defourny, P., Leeuw, G.d., Forsberg, R., Holzer-Popp, T., Paul, F., Sandven, S., Sathyendranath, S., Roozendaal, M.v., Wagner, W., 2013. The ESA Climate Change Initiative: satellite data records for essential climate variables. *Bull. Am. Meteorol. Soc.* 94, 1541–1552.
- Homer, C.C., Huang, L., Yang, B., Wylie, B., Coan, M., 2004. Development of a 2001 national land-cover database for the United States. *Photogramm. Eng. Remote Sens.* 70, 829–840.
- Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., Herold, N., Mckerrow, A., Van Driel, J.N., Wickham, J., 2007. Completion of the 2001 National Land Cover Database for the conterminous United States. *Photogramm. Eng. Remote Sens.* 73, 337–341.
- Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., Megown, K., 2015. Completion of the 2011 National Land Cover Database for the conterminous United States-representing a decade of land cover change information. *Photogramm. Eng. Remote Sens.* 81, 345–354.

- Huang, C., Coward, S.N., Masek, J.G., Thomas, N., Zhu, Z., Vogelmann, J.E., 2010. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sens. Environ.* 114, 183–198.
- Jakubauskas, M.E., Lulla, K.P., Mausell, P.W., 1990. Assessment of vegetation change in a fire-altered forest landscape. *Photogramm. Eng. Remote. Sens.* 56, 371–377.
- Kasischke, E.S., French, N.H.F., 1995. Locating and estimating the areal extent of wildfires in Alaskan Boreal Forests using multiple-season AVHRR NDVI composite data. *Remote Sens. Environ.* 51, 263–275.
- Kennedy, R.E., Cohen, W.B., 2003. Automated designation of tie-points for image-to-image coregistration. *Int. J. Remote Sens.* 24, 3467–3490.
- Kennedy, R.E., Yang, Z., Cohen, W.B., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr – temporal segmentation algorithms. *Remote Sens. Environ.* 114, 2897–2910.
- Kennedy, R.E., Andrefouet, S., Cohen, W.B., Gomez, C., Griffiths, P., Hais, M., Healey, S.P., Helmer, E.H., Hostert, P., Lyons, M.B., Meigs, G.W., Pflugmacher, D., Phinn, S.R., Powell, S.L., Scarth, P., Sen, S., Schroeder, T.A., Schneider, A., Sonnenschein, R., Vogelmann, J.E., Wulder, M.A., Zhu, Z., 2014. Bringing an ecological view of change to Landsat-based remote sensing. *Front. Ecol. Environ.* 12, 339–346.
- Kennedy, R.E., Yang, Z.Q., Braaten, J., Copass, C., Antonova, N., Jordan, C., Nelson, P., 2015. Attribution of disturbance change agent from Landsat time-series in support of habitat monitoring in the Puget Sound region, USA. *Remote Sens. Environ.* 166, 271–285.
- Key, C.H., Benson, N.C., 2006. Landscape assessment: remote sensing of severity, the Normalized Burn Ratio. In: Lutes, D.C., et al. (Eds.), FIREMON: Fire Effects Monitoring and Inventory System. USDA Forest Service, Rocky Mountain Research Station, General Technical Report, RMRS-GTR-164-CD:LA1-LA51, Ogden, UT.
- Koutsias, N., 2003. An autologistic regression model for increasing the accuracy of burned surface mapping using Landsat Thematic Mapper data. *Int. J. Remote Sens.* 24, 2199.
- Koutsias, N., Karteris, M., 1998. Logistic regression modelling of multitemporal Thematic Mapper data for burned area mapping. *Int. J. Remote Sens.* 19, 3499–3514.
- Koutsias, N., Karteris, M., 2000. Burned area mapping using logistic regression modeling of a single post-fire Landsat-5 Thematic Mapper image. *Int. J. Remote Sens.* 21, 673–687.
- Kovalskyy, V., Roy, D.P., 2013. The global availability of Landsat 5 TM and Landsat 7 ETM+ land surface observations and implications for global 30 m Landsat data product generation. *Remote Sens. Environ.* 130, 280–293.
- Krawchuk, M.A., Moritz, M.A., Parisien, M.A., Van Dorn, J., Hayhoe, K., 2009. Global pyrogeography: the current and future distribution of wildfire. *PLoS One* 4.
- Kushla, J.D., Ripple, W.J., 1998. Assessing wildfire effects with Landsat thematic mapper data. *Int. J. Remote Sens.* 19, 2493.
- Li, Z.Q., Cihlar, J., Moreau, L., Huang, F.T., Lee, B., 1997. Monitoring fire activities in the boreal ecosystem. *J. Geophys. Res.-Atmos.* 102, 29611–29624.
- Liang, L., Chen, Y., Hawbaker, T.J., Zhu, Z., Gong, P., 2014. Mapping mountain pine beetle mortality through growth trend analysis of time-series Landsat data. *Remote Sens.* 6, 5696–5716.
- Liu, Z.H., 2016. Effects of climate and fire on short-term vegetation recovery in the boreal larch forests of Northeastern China. *Sci Rep* 6.
- López García, M.J., Caselles, V., 1991. Mapping burns and natural reforestation using thematic Mapper data. *Geocarto Int.* 6 (1):31–37. <http://www.tandfonline.com/doi/abs/10.1080/10106049109354290>.
- Malingreau, J.P., Stephens, G., Fellows, L., 1985. Remote sensing of forest fires: Kalimantan and North Borneo in 1982–83. *Ambio* 14, 314–321.
- Markham, B.L., Helder, D.L., 2012. Forty-year calibrated record of earth-reflected radiance from Landsat: a review. *Remote Sens. Environ.* 122, 30–40.
- Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., Feng, G., Kutler, J., Teng-Kui, L., 2006. A Landsat surface reflectance dataset for North America, 1990–2000. *IEEE Geosci. Remote Sens. Lett.* 3, 68–72.
- Matson, M., Dozier, J., 1981. Identification of subresolution high-temperature sources using a thermal IR sensor. *Photogramm. Eng. Remote. Sens.* 47, 1311–1318.
- Matson, M., Stephens, G., Robinson, J., 1987. Fire detection using data from the NOAA-N satellites. *Int. J. Remote Sens.* 8, 961–970.
- McCarty, J.L., Korontzi, S., Justice, C.O., Loboda, T., 2009. The spatial and temporal distribution of crop residue burning in the contiguous United States. *Sci. Total Environ.* 407, 5701–5712.
- McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* 17, 1425–1432.
- Meddens, A.J.H., Kolden, C.A., Lutz, J.A., 2016. Detecting unburned areas within wildfire perimeters using Landsat and ancillary data across the northwestern United States. *Remote Sens. Environ.* 186:275–285. <http://dx.doi.org/10.1016/j.rse.2016.08.023>.
- Merentitis, A., Debes, C., 2015. Many hands make light work-on ensemble learning techniques for data fusion in remote sensing. *IEEE Geosci. Remote Sens. Lett.* Mag. 3, 86–99.
- Michalek, J.L., French, N.H.F., Kasischke, E.S., Johnson, R.D., Colwell, J.E., 2000. Using Landsat TM data to estimate carbon release from burned biomass in an Alaskan spruce forest complex. *Int. J. Remote Sens.* 21, 323.
- Miller, J.D., Yool, S.R., 2002. Mapping forest post-fire canopy consumption in several over-story types using multi-temporal Landsat TM and ETM data. *Remote Sens. Environ.* 82, 481.
- Millman, K.J., Aivazis, M., 2011. Python for scientists and engineers. *Comput. Sci. Eng.* 13, 9–12.
- Morisette, J.T., Giglio, L., Csizsar, I., Setzer, A., Schroeder, W., Morton, D., Justice, C.O., 2005. Validation of MODIS active fire detection products derived from two algorithms. *Earth Interact.* 9, 9–25.
- Morisette, J.T., Baret, F., Liang, S., 2006. Special issue on global land product validation. *IEEE Trans. Geosci. Remote Sens.* 44, 1695–1697.
- National Research Council, 2001. Transforming Remote Sensing Data into Information and Applications. Available at: <http://www.nap.edu/catalog/10257/transforming-remote-sensing-data-into-information-and-applications> The National Academies Press, Washington, DC.
- Oliphant, T.E., 2007. Python for scientific computing. *Comput. Sci. Eng.* 9, 10–20.
- Omernik, J.M., Griffith, G.E., 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. *Environ. Manag.* 54, 1249–1266.
- Open Source Geospatial Foundation, 2016. GDAL - Geospatial Data Abstraction Library: Version 1.11.2.
- Padilla, M., Stehman, S.V., Ramo, R., Corti, D., Hantson, S., Oliva, P., Alonso-Canas, I., Bradley, A.V., Tansey, K., Mota, B., Pereira, J.M., Chuvieco, E., 2015. Comparing the accuracies of remote sensing global burned area products using stratified random sampling and estimation. *Remote Sens. Environ.* 160, 114–121.
- Parisien, M.A., Moritz, M.A., 2009. Environmental controls on the distribution of wildfire at multiple spatial scales. *Ecol. Monogr.* 79, 127–154.
- Patterson, M.W., Yool, S.R., 1998. Mapping fire-induced vegetation mortality using landsat thematic mapper data: a comparison of linear transformation techniques. *Remote Sens. Environ.* 65, 132–142.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Courapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Powell, S.L., Cohen, W.B., Healey, S.P., Kennedy, R.E., Moisen, G.G., Pierce, K.B., Ohmann, J.L., 2010. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: a comparison of empirical modeling approaches. *Remote Sens. Environ.* 114, 1053–1068.
- Prins, E.M., Menzel, W.P., 1992. Geostationary satellite detection of bio mass burning in South America. *Int. J. Remote Sens.* 13, 2783–2799.
- Prins, E.M., Menzel, W.P., 1994. Trends in South American biomass burning detected with the GOES visible infrared spin scan radiometer atmospheric sounder from 1983 to 1991. *J. Geophys. Res.-Atmos.* 99, 16719–16735.
- Python Software Foundation, 2016. Python Language Reference, Version 2.7. Available at <http://www.python.org>.
- R Foundation for Statistical Computing, 2014. R: A Language and Environment for Statistical Computing, v. 3.1.1 (Vienna, Austria).
- Randerson, J.T., Chen, Y., van der Werf, G.R., Rogers, B.M., Morton, D.C., 2012. Global burned area and biomass burning emissions from small fires. *J. Geophys. Res. Biogeosci.* 117 (n/a-n/a). (doi:10.1029/2012JG002128).
- Riley, K.L., Abatzoglou, J.T., Grenfell, I.C., Klene, A.E., Heinsch, F.A., 2013. The relationship of large fire occurrence with drought and fire danger indices in the western USA, 1984–2008: the role of temporal scale. *Int. J. Wildland Fire* 22, 894–909.
- Rogan, J., Yool, S.R., 2001. Mapping fire-induced vegetation depletion in the Peloncillo Mountains, Arizona and New Mexico. *Int. J. Remote Sens.* 22, 3101–3121.
- Rogan, J., Franklin, J., Roberts, D.A., 2002. A comparison of methods for monitoring multitemporal vegetation change using Thematic Mapper imagery. *Remote Sens. Environ.* 80, 143–156.
- Roy, D.P., Jin, Y., Lewis, P.E., Justice, C.O., 2005. Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data. *Remote Sens. Environ.* 97, 137.
- Roy, D.P., Wulder, M.A., Loveland, T.R., CE, W., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., Scambos, T.A., Schaaf, C.B., Schott, J.R., Sheng, Y., Vermote, E.F., Belward, A.S., Bindschadler, R., Cohen, W.B., Gao, F., Hipple, J.D., Hostert, P., Huntington, J., Justice, C.O., Kilic, A., Kovalskyy, V., Lee, Z.P., Lymburner, L., Masek, J.G., McCorkel, J., Shuai, Y., Trezza, R., Vogelmann, J., Wynne, R.H., Zhu, Z., 2014. Landsat-8: Science and product vision for terrestrial global change research. *Geocarto Int.* 145:154–172. <http://dx.doi.org/10.1016/j.jrse.2014.02.001>.
- Ryan, K.C., Knapp, E.E., Varner, J.M., 2013. Prescribed fire in North American forests and woodlands: history, current practice, and challenges. *Front. Ecol. Environ.* 11, e15–e24.
- Schneider, A., 2012. Monitoring land cover change in urban and pen-urban areas using dense time stacks of Landsat satellite data and a data mining approach. *Remote Sens. Environ.* 124, 689–704.
- Schroeder, W., Prins, E., Giglio, L., Csizsar, I., Schmidt, C., Morisette, J., Morton, D., 2008. Validation of GOES and MODIS active fire detection products using ASTER and ETM+ data. *Remote Sens. Environ.* 112, 2711–2726.
- Schroeder, W., Oliva, P., Giglio, L., Quayle, B., Lorenz, E., Morelli, F., 2015. Active fire detection using Landsat-8/OLI data. *Remote Sens. Environ.* 185, 210–220.
- Shakesby, R.A., Doerr, S.H., 2006. Wildfire as a hydrological and geomorphological agent. *Earth Sci. Rev.* 74, 269–307.
- Short, K.C., 2014. A spatial database of wildfires in the United States, 1992–2011. *Earth Syst. Sci. Data* 6, 1–27.
- Short, K.C., 2015. Sources and implications of bias and uncertainty in a century of US wildfire activity data. *Int. J. Wildland Fire* 24, 883–891.
- Stroppiana, D., Bordogna, G., Carrara, P., Boschetti, M., Boschetti, L., Brivio, P.A., 2012. A method for extracting burned areas from Landsat TM/ETM+ images by soft aggregation of multiple Spectral Indices and a region growing algorithm. *ISPRS J. Photogramm. Remote Sens.* 69, 88–102.
- Thomas, N.E., Huang, C., Goward, S.N., Powell, S., Rishmawi, K., Schleeweis, K., Hinds, A., 2011. Validation of North American Forest Disturbance dynamics derived from Landsat time series stacks. *Remote Sens. Environ.* 115 (1):19–32. <http://www.sciencedirect.com/science/article/pii/S0034425710002257>.
- Thompson, D.K., Simpson, B.N., Beaudoin, A., 2016. Using forest structure to predict the distribution of tree boreal peatlands in Canada. *For. Ecol. Manag.* 372, 19–27.
- Trigg, S., Flasse, S., 2000. Characterizing the spectral-temporal response of burned savannah using in situ spectroradiometry and infrared thermometry. *Int. J. Remote Sens.* 21, 3161–3168.
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150.

- van der Walt, S., Colbert, S.C., Varoquaux, G., 2011. The NumPy array: a structure for efficient numerical computation. *Comput. Sci. Eng.* 13, 22–30.
- van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., Yu, T., the scikit-image contributors, 2014. *scikit-image: image processing in Python*. *PeerJ* 2, e453.
- van der Werf, G.R., Randerson, J.T., Giglio, L., Collatz, G.J., Mu, M., Kasibhatla, P.S., Morton, D.C., DeFries, R.S., Jin, Y., van Leeuwen, T.T., 2010. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmos. Chem. Phys.* 10, 11707–11735.
- Vanderhoof, M.K., Burnner, N.M., Beal, Y., Hawbaker, T.J., 2017. Validation of the Landsat burned area ECV across the conterminous U.S. using a semi-automated burned area reference dataset. *Remote Sens. Environ.* 198, 393–406.
- Verbesselt, J., Hyndman, R., Newnham, G., Culvenor, D., 2010. Detecting trend and seasonal changes in satellite image time series. *Remote Sens. Environ.* 114, 106–115.
- Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K., Van Driel, J.N., 2001. Completion of the 1990's National Land Cover Data Set for the conterminous United States. *Photogramm. Eng. Remote. Sens.* 67, 650–662.
- Westerling, A.L., Hidalgo, H.G., Cayan, D.R., Swetnam, T.W., 2006. Warming and earlier spring increase western US forest wildfire activity. *Science* 313, 940–943.
- Westerling, A.L., Turner, M.G., Smithwick, E.A.H., Romme, W.H., Ryan, M.G., 2011. Continued warming could transform Greater Yellowstone fire regimes by mid-21st century. *Proc. Natl. Acad. Sci. U. S. A.* 108, 13165–13170.
- Williams, C.A., Gu, H., MacLean, R., Masek, J.G., Collatz, G.J., 2016. Disturbance and the carbon balance of US forests: a quantitative review of impacts from harvests, fires, insects, and droughts. *Glob. Planet. Chang.* 143, 66–80.
- Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2012. Opening the archive: how free data has enabled the science and monitoring promise of Landsat. *Remote Sens. Environ.* 122, 2–10.
- Zhao, F., Huang, C., Zhu, Z., 2015. Use of vegetation change tracker and support vector machine to map disturbance types in Greater Yellowstone Ecosystems in a 1984–2010 Landsat time series. *IEEE Geosci. Remote Sens. Lett.* 12, 1650–1654.
- Zhu, Z., Woodcock, C.E., 2014a. Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: an algorithm designed specifically for monitoring land cover change. *Remote Sens. Environ.* 152, 217–234.
- Zhu, Z., Woodcock, C.E., 2014b. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sens. Environ.* 144, 152–171.