Remote Sensing of Environment 198 (2017) 393–406

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Validation of the USGS Landsat Burned Area Essential Climate Variable (BAECV) across the conterminous United States

Melanie K. Vanderhoof *, Nicole Fairaux, Yen-Ju G. Beal, Todd J. Hawbaker

U.S. Geological Survey, Geosciences and Environmental Change Science Center, P.O. Box 25046, DFC, MS980, Denver, CO 80225, United States

ARTICLE INFO

Article history:
Received 2 November 2016
Received in revised form 6 June 2017
Accepted 25 June 2017
Available online xxxx

Keywords:
Burned area
Essential climate variable
Fire
Landsat
Validation
Wildfire

ABSTRACT

The Landsat Burned Area Essential Climate Variable (BAECV), developed by the U.S. Geological Survey (USGS), capitalizes on the long temporal availability of Landsat imagery to identify burned areas across the conterminous United States (CONUS) (1984–2015). Adequate validation of such products is critical for their proper usage and interpretation. Validation of coarse-resolution products often relies on independent data derived from moderate-resolution sensors (e.g., Landsat). Validation of Landsat products, in turn, is challenging because there is no corresponding source of high-resolution, multispectral imagery that has been systematically collected in space and time over the entire temporal extent of the Landsat archive. Because of this, comparison between high-resolution images and Landsat science products can help increase user's confidence in the Landsat science products, but may not, alone, be adequate. In this paper, we demonstrate an approach to systematically validate the Landsat-derived BAECV product. Burned area extent was mapped for Landsat image pairs using a manually trained semi-automated algorithm that was manually edited across 28 path/rows and five different years (1988, 1993, 1998, 2003, 2008). Three datasets were independently developed by three analysts and the datasets were integrated on a pixel by pixel basis in which at least one to all three analysts were required to agree a pixel was burned. We found that errors within our Landsat reference dataset could be minimized by using the rendition of the dataset in which pixels were mapped as burned if at least two of the three analysts agreed. BAECV errors of omission and commission for the detection of burned pixels averaged 42% and 33%, respectively for CONUS across all five validation years. Errors of omission and commission were lowest across the western CONUS, for example in the shrub and scrublands of the Arid West (31% and 24%, respectively), and highest in the grasslands and agricultural lands of the Great Plains in central CONUS (62% and 57%, respectively). The BAECV product detected most (>65%) fire events >10 ha across the western CONUS (Arid and Mountain West ecoregions). Our approach and results demonstrate that a thorough validation of Landsat science products can be completed with independent Landsat-derived reference data, but could be strengthened by the use of complementary sources of high-resolution data.

Published by Elsevier Inc.

1. Introduction

Accurate mapping of the extent and timing of burned area is critical to quantifying and modeling greenhouse gas emissions (Crutzen and Andreae, 1990; Palacios-Orueta et al., 2005; Randerson et al., 2005), carbon and nutrient cycling (Conard et al., 2002; Bond-Lamberty et al., 2007), and changes to ecosystem structure (Thonicke et al., 2001; Goetz et al., 2005). Consequently, fire disturbance has been identified by the Global Climate Observing System (GCOS) program as one of the high priority Essential Climate Variables (ECV) (Global Climate Observing System, 2004) and major efforts have been undertaken to produce global burned area products (Mouillot et al., 2014). The products developed to-date use coarse-scale satellite imagery (300 m to 1 km) (e.g., Moderate Resolution Imaging Spectrometer (MODIS) burned area product (MCD45, MCD64), Geoland2, fire_cci burned area (BA)). Such datasets provide information critical for climate modeling and are effective for capturing global fire patterns at a high temporal frequency, but may be limited in their ability to map fire heterogeneity, detect small fires (Stroppiana et al., 2012) or provide enough historical context, necessary to discern temporal trends (Mouillot et al., 2014) and relationships with climate and other drivers (Podur et al., 2002; Miller et al., 2009; Whitman et al., 2015). In addition, because of the tremendous amount of spectral diversity in the signal of burned areas across diverse vegetation types, fire combustion levels (e.g., ash, char, soot), and burn severities (e.g., ground vs crown fires), the accuracy of existing global burned area products is relatively low with documented errors of omission and commission for burned areas ranging from 51%

* Corresponding author.
E-mail address: mvanderhoof@usgs.gov (M.K. Vanderhoof).

http://dx.doi.org/10.1016/j.rse.2017.06.025
0034-4257/Published by Elsevier Inc.

to 93%, and 42% to 94%, respectively (Padilla et al., 2014a, 2015; Chuvieco et al., 2016).

This study presents a validation of the Landsat Burned Area Essential Climate Variable (BAECV) product across a sample set of locations and times using an independently derived reference dataset. The BAECV, developed by the U.S. Geological Survey (USGS), aims to capitalize on the long time period covered by Landsat imagery to provide wall-to-wall maps of burned areas across the conterminous United States (CONUS) (1984–2015), and could be extended to other regions with appropriate training data (Hawbaker et al., 2017). The product will be provided as a wall-to-wall raster of burned area across CONUS at 30 m resolution and an annual time-step, with a minimum fire size of 4.05 ha (45 pixels) (https://www.sciencebase.gov/catalog/item/57867943e4b0e02680c14fec). Landsat sensors can provide a longer temporal record of burned area relative to existing global burned area products and potentially with increased accuracy and detail (Stroppiana et al., 2012). Landsat has been used extensively to map burned areas, predominantly for local and regional studies (Mitri and Gitas, 2004; Bastarrika et al., 2011; Petropoulos et al., 2011; Mallinis and Koutsias, 2012). In recent years Landsat has been used to map fire and other disturbance types across portions of CONUS but for limited years (Masek et al., 2008; Boschetti et al., 2015) and/or Landsat path/rows (Cohen et al., 2010; Kennedy et al., 2010; Thomas et al., 2011). These efforts have been largely restricted to forest cover and fire is often not distinguished from other disturbance types (e.g., harvest, insect) (Goward et al., 2016). The most comparable effort to date is the Monitoring Trends in Burn Severity (MTBS) product which is also derived from Landsat (Eidenshink et al., 2007). The BAECV differs from the MTBS dataset in several important ways. Because the BAECV product generation is automated, the BAECV can potentially provide a more complete census of burned areas, relative to the MTBS dataset which relies on manually mapping reported large fires ($\geq$2 km^2 in the eastern U.S. and $\geq$4 km^2 in the western U.S.) (Eidenshink et al., 2007). The BAECV utilizes all available Landsat images, in contrast the MTBS effort began prior to the Landsat archive becoming freely available in 2008, which required them to be strategic in their image selection for the earlier years of the dataset. In addition, the MTBS made a conscious decision to provide limited mapping of prescribed fires, common in the southeastern United States due to the sheer number of such fires. Despite the advantages, automation can be expected to introduce errors in burned area extent (e.g., missing fires, over-mapping fires, or disagreeing on fire extent), necessitating an independent validation of the BAECV product.

Validation of burned area products and the provision of accuracy statistics to users is essential to allow users to decide when and how to utilize datasets, correctly interpret results, and provide feedback to improve products (Roy et al., 2005; Morisette et al., 2006). The Committee on Earth Observation Satellites (CEOS), Land Product Validation Subgroup (LPVS), formed in 2000, has specified that validation is a critical component in the generation of ECV products, and should follow internationally agreed upon validation best practices to measure accuracy, precision (standard error of accuracy estimates), and temporal stability at comprehensive spatial and temporal scales (Morisette et al., 2006). Validations typically produce pixel-level or point-level error matrices, derived by cross-tabulating ECV products with independent reference maps (Bastarrika et al., 2011; Stroppiana et al., 2012; Padilla et al., 2014a, 2015). Linear regression analysis has also been used to compare the proportion of burned area defined by the product, relative to reference maps (Roy et al., 2008; Roy and Boschetti, 2009). Comparisons between global burned area products have also examined differences in the spatial and temporal distribution of burned area and calculated patch indices to explore a product's ability to map small fires (Chuvieco et al., 2016).

The source of reference datasets varies by study and product. Fire perimeter datasets, such as the U.S. Geospatial Multi-Agency Coordination (GeoMAC) dataset, tend to either focus on large fires or are designed to meet the needs of fire managers and do not provide a complete census of all fires (Eidenshink et al., 2007; Walters et al., 2011). This design makes perimeter datasets good references for large, single fire events (Mitri and Gitas, 2004; Henry, 2008; Bastarrika et al., 2011), but insufficient for a validation at a national or global scale. Because of the limitations of fire perimeter datasets, burned area maps derived from remotely sensed imagery are typically validated using a reference map derived from a finer resolution source of imagery (Roy and Boschetti, 2009; Mallinis and Koutsias, 2012; Padilla et al., 2014a). For coarse-resolution products this is non-problematic, as moderate-resolution sources of imagery, collected at regular intervals, are widely available (e.g., Landsat and ASTER) (Roy and Boschetti, 2009).

The utilization of high-resolution imagery (e.g., IKONOS, Quickbird-2, Geoeye-1, Worldview-2, 3) to validate a national or global burned area product, however, faces several challenges. High-resolution imagery has been successfully utilized to detect burned areas (Mitri and Gitas, 2006; Holden et al., 2010; Mallinis and Koutsias, 2012). Yet, as the satellites typically collect imagery on demand, the coincidence of images collected over burned patches, prior to vegetation recovery, is sporadic making it challenging to defend a sampling strategy and requiring classification of burned area from a single image instead of a pre- and post-fire image pair. In addition, these satellites have only been in orbit since late 1990s or early 2000s, meaning they can only be used to validate a portion of the temporal extent of a Landsat science product. These satellites also typically lack short-wave infrared (SWIR) bands, which have been found to be useful in detecting burned areas (Chuvieco, 1997). The spatial scale at which fire events occur should also be considered relative to the reference data. The small extent of high-resolution images (13 to 18 km across), relative to Landsat image extents (185 km across) means only portions of larger fires are often contained within high-resolution imagery, limiting the number of fire events being validated.

Instead of using high-resolution imagery, validation of Landsat disturbance products to date have typically relied on the derivation of independent datasets from Landsat images, complemented by high-resolution imagery, as available (Thomas et al., 2011; Masek et al., 2013). Burned patches are often visually distinct, but have high spectral diversity resulting from variability in soil type, pre-fire vegetation cover, fire severity and time since fire, and that can make it challenging to detect burned areas across diverse environments in an automated manner (Bastarrika et al., 2011). Therefore, forest disturbance events are identified through the visual examination of pre- and post- Landsat images by experienced image analysts (Masek et al., 2008; Huang et al., 2009; Stroppiana et al., 2012). Although including a manual component in imagery analysis is a common practice to improve the quality of reference datasets (e.g., Mitri and Gitas, 2004; Henry, 2008; Petropoulos et al., 2011), observer-dependent variability has also been documented, although not explicitly for mapping burned areas (Mazz, 1996; Baveye et al., 2010). Using multiple observers is one technique that has been used to reduce errors of omission in other areas of science and image analysis, but is not widely done (Mazz, 1996; Nichols et al., 2000).

Thorough validation of remote sensing products is essential prior to their acceptance by the scientific community, proper use, and integration into management and modeling activities. This study seeks to validate USGS's Landsat BAECV (1984–2015) using an independent dataset derived from Landsat across a sample of 28 Thiessen scene areas and five years, complemented by high-resolution imagery. Our research questions included:

(1). How does the subjectivity of visual image interpretation affect the quality of the reference dataset and influence accuracy statistics?

(2). What is the accuracy of USGS's BAECV across diverse land cover types and regions of the conterminous U.S. (CONUS)?

(3). How stable are the accuracy statistics through time?

(4). How does burn size influence the accuracy of the BAECV product?