ARTICLE IN PRESS
RSE-10325; No of Pages 16
Remote Sensing of Environment xxx (2017) xxx–xxx
Contents lists available at ScienceDirect
Remote Sensing of Environment
journal homepage: www.elsevier.com/locate/rse

# Stratification and sample allocation for reference burned area data

Marc Padilla [a,*], Pontus Olofsson [b], Stephen V. Stehman [c], Kevin Tansey [a], Emilio Chuvieco [d]

[a] Centre for Landscape and Climate Research, Department of Geography, University of Leicester, UK
[b] Department of Earth and Environment, Boston University, USA
[c] Department of Forest and Natural Resources Management, College of Environmental Science and Forestry, State University of New York, Syracuse, NY, 13210, USA
[d] Environmental Remote Sensing Research Group, Department of Geology, Geography and Environment, Universidad de Alcalá, Spain

ARTICLE INFO

*Article history:*
Received 30 September 2016
Received in revised form 26 June 2017
Accepted 29 June 2017
Available online xxxx

*Keywords:*
Validation
Probability sampling
Terrestrial globe
Fire Disturbance
Sampling efficiency

ABSTRACT

Statistical estimation protocols are one of the key means to ensure that independent and objective information on product accuracy is communicated to end-users. Methods for validating burned area products have been developed based on a probability sample of a space by time partitioning of the population. We extend this basic methodology to improve stratification and sample allocation, key elements of a sampling design used to collect burned area reference data. We developed and evaluated an approach to partition each year and biome into low and high burned area (*BA*) strata. Because the threshold used to separate the sampling units into low and high *BA* can vary by year and biome, this approach offers a more targeted stratification than used in previous studies for which a common threshold was applied to all biomes. A hypothetical population of validation data was then used to quantitatively compare the precision of accuracy estimates derived from different stratification and sample size allocation options. We evaluated two options that had been previously examined in the *BA* validation literature, and extended previous studies by adding two new options specifically developed for ratio estimates. Stratification based on mapped *BA* reduced standard errors of the global burned area accuracy estimates from one-half to one-eighth relative to standard errors of simple random sampling. Stratifying by mapped *BA* was also found to reduce standard errors of accuracy estimates for most year by biome strata indicating that this advantage of stratification and sample allocation applies generally to a range of conditions (i.e., biomes and years). The most precise estimates were obtained using a sample size per stratum allocation $n_h \propto N_h \sqrt{\overline{BA}_h}$ where $N_h$ is the number of units in stratum $h$ and $\overline{BA}_h$ is the mean mapped *BA* for stratum $h$. The best sampling design from our analyses was then used to select a set of 1,000 samples from a hypothetical population of validation data and confidence intervals were computed for each sample. Close to 95% of these confidence intervals contained the true population value thus confirming the validity of confidence intervals produced from the estimates and standard errors.



## 1. Introduction

Biomass burning is one of the most important processes impacting the Earth system (Bond and Keeley, 2005; Bowman et al., 2009) and one of the main sources of gases and aerosols emitted to the atmosphere (van der Werf et al., 2004, 2010). The Global Climate Observing System (GCOS) program identified Fire Disturbance as an Essential Climate Variable (ECV) (GCOS, 2004), commonly expressed by burned area (*BA*) information (Mouillot et al., 2014). Global *BA* products provide the location and dates of burned surfaces at a coarse spatial resolution (300–1000 m).

Product validation is defined as "… the process of assessing, by independent means, the quality of the data products derived from the system outputs" (CEOS-WGCV, 2012). *BA* products usually cover multi-year periods and the Committee on Earth Observation Satellites (CEOS) Land Product Validation Subgroup (LPV) highlights the importance of assessing the temporal stability of a product's accuracy by collecting data over globally representative locations and time periods (http://lpvs.gsfc.nasa.gov). The selection of representative samples is particularly important when the event to be characterized is rare and occurs in spatio-temporal clusters, e.g. fires (Chou et al., 1993; Giglio et al., 2010). The main challenge is to define an optimal sampling design that leads to precise accuracy estimates and allocates the sample through several time periods and regions of interest, e.g. years in a multiyear time period and major biomes. Throughout the manuscript optimal sampling design refers to the design that minimizes the variance of an accuracy estimate, for a specific sample size (Cochran, 1977; Section 5.5). In our application, the factors evaluated that can affect the optimal design are the strata and the allocation of the sample to these strata.

* Corresponding author.
*E-mail address:* mp489@le.ac.uk (M. Padilla).

http://dx.doi.org/10.1016/j.rse.2017.06.041
0034-4257/

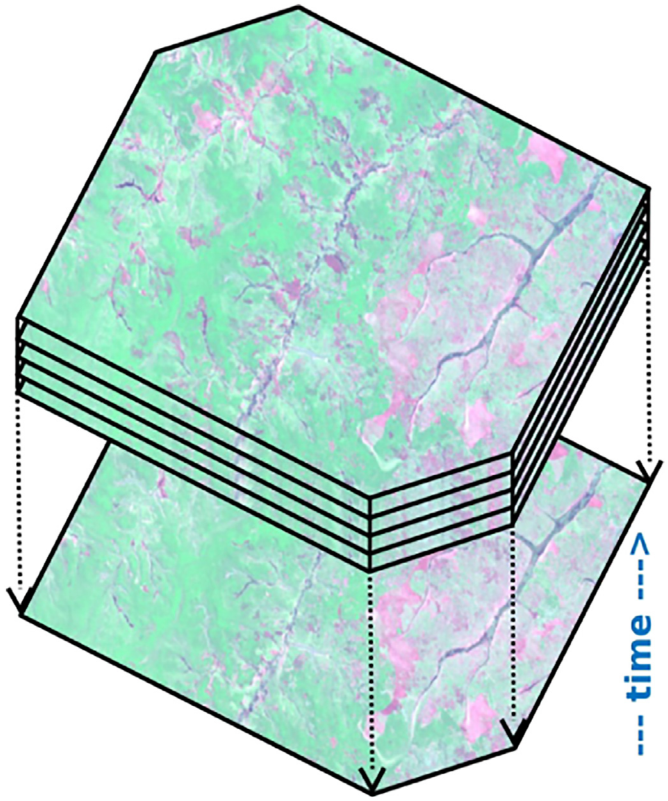


**Fig. 1.** Illustration of the "voxel" sampling units proposed by Boschetti et al. (2016) for partitioning the three-dimensional space by time population. Each sampling unit is delimited spatially by a Thiessen Scene Area (TSA) partitioning the two dimensions of space and temporally (the third dimension) by the time between two consecutive Landsat images. The image at the bottom is displaced further down to illustrate temporal spectral changes. Images are displayed as false color composites with SWIR, NIR and red bands in the red, green and blue channels respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

*BA* reference data are commonly obtained following the recommendations of the CEOS LPV (Boschetti et al., 2009), using multi-temporal pairs of medium spatial resolution images (10–60 m). Methodological limitations and reference data generation costs led until recently to relatively small sample sizes, the selection of sites being based on expert knowledge (Chuvieco et al., 2008; Giglio et al., 2009; Padilla et al., 2014b; Plummer et al., 2007; Roy and Boschetti, 2009; Tansey et al., 2008). Recently, global accuracy estimates were first produced from a probability sampling design consisting of a spatially stratified random sample (Padilla et al., 2014a, 2015). Boschetti et al. (2016) developed a sampling approach based on partitioning the space by time population into three-dimensional "voxel" units defined by Thiessen Scene Areas (TSAs) to partition space and 16-day Landsat image pairs to partition time. Although the specific example presented by Boschetti et al. (2016) used TSAs and a time period specific to Landsat, the general approach using the voxel units can be applied to other spatial and temporal partitions. Boschetti et al. (2016) also proposed a stratification based on a threshold of active fire counts that split each biome into two strata representing low and high fire activity. For all geographic strata (biomes), Boschetti et al. (2016) determined this threshold using the 20th percentile of the cumulative distribution of active fire counts. Boschetti et al. (2016) acknowledged that additional work was needed to investigate "the impact of different thresholds to define low and high fire activity." Furthermore, the allocation methods they analysed did not include two methods recommended by Cochran (1977; Section 6.14) for ratio estimates. The main accuracy measures for burned area are in fact ratios, as is for example the case of the commission and omission error rates.

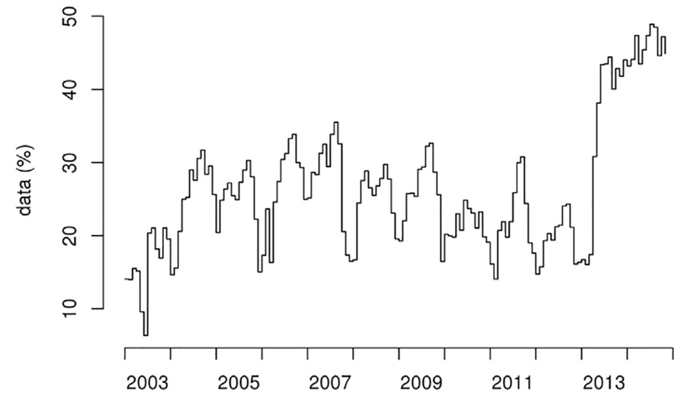


**Fig. 3.** Temporal distribution of reference data availability expressed as monthly percentage of area·time covered by Landsat image pairs separated by 16 days or fewer.

The objective of the current study is to improve the sampling design by: (a) allowing the threshold used to define low and high burned area strata to vary by biome and year, and (b) identifying the best method for allocating the sample to strata among those methods recommended in the literature. The precision of the accuracy estimates obtained for the different stratification and sample allocation methods was compared using a hypothetical population of validation data. The hypothetical population allows for direct comparison of the standard errors of accuracy estimates for the different design options evaluated. These comparisons provide quantitative information to guide decisions regarding how to construct strata separating low and high burned area and how to effectively allocate the sample to these strata.

The current research is in the framework of the validation effort of the Fire Disturbance (Fire_cci) Phase II project (www.esa-fire-cci.org), part of the European Space Agency's (ESA) Climate Change Initiative. The main goal of this effort is to generate reference data that cover the twelve year period 2003–2014. This period was defined by the time period for which data were available for both global sensors used in the project, MEdium Resolution Imaging Spectrometer (MERIS) and Moderate Resolution Imaging Spectroradiometer (MODIS).

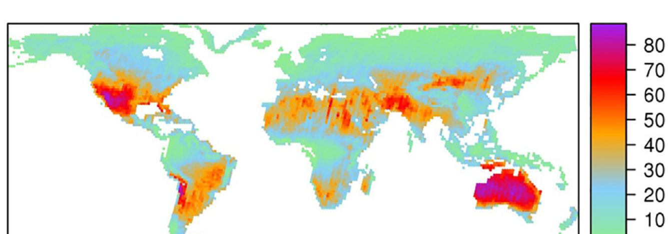


**Fig. 2.** Spatial distribution of reference data availability for 2003 to 2014 expressed as percentage of time that Thiessen Scene Areas covered by Landsat image pairs separated by 16 days or fewer are available from the USGS archive (accessed on September 2015).

**Table 1**
Population error matrix for a single Thiessen Scene Area (TSA) sampling unit. Matrix cells express agreement (diagonal cells) or disagreement (off-diagonal cells) in terms of area ($m^2$) between the *BA* product (map) class and the reference classification of the entire unit.

| Product classification | Reference classification | | Row total |
|---|---|---|---|
| | Burned | Unburned | |
| Burned | $E_{11}$ | $E_{12}$ | $E_{1+}$ |
| Unburned | $E_{21}$ | $E_{22}$ | $E_{2+}$ |
| Col. Total | $E_{+1}$ | $E_{+2}$ | |

Please cite this article as: Padilla, M., et al., Stratification and sample allocation for reference burned area data, Remote Sensing of Environment (2017), http://dx.doi.org/10.1016/j.rse.2017.06.041