

## ***Interactive comment on “Development of a standard database of reference sites for validating global burned area products” by Magí Franquesa et al.***

### **Anonymous Referee #1**

Received and published: 21 June 2020

#### General

The reference datasets for validating global burned area products provide a valuable resource to the fire mapping community. As the authors note, collecting reference data to validate burned area products is an expensive and time consuming proposition. Having available a vetted set of reference sample sites for map producers to readily access will greatly enhance the quantity and quality of information available to assess and compare accuracy of burned area products. The global extent of these datasets will facilitate regional comparisons as well, as users of the data will be able to extract data specific to their study area. One of the fundamental challenges of mapping of any

C1

theme, burned area or otherwise, is the immense difficulty of obtaining reference data. The burned area reference dataset (BARD) presented by the authors is a significant advance to diminish this difficulty.

#### Specific Comments

1. The authors astutely identify the role of sampling in the collection of these burned area reference datasets (Line 63). It would be useful to add some explanation distinguishing between reference data collected by a formal sampling design, often called probability sampling designs, and reference data collected by convenience, ease of access, or other method that does not necessarily have randomization. Reference data collected by a randomized sampling design are suitable to support rigorous statistical statements about accuracy, whereas data collected by convenience can be suspect in this regard (i.e., data may not be representative of the entire area of interest). The implications of how the reference data were obtained should be noted. The manuscript clearly indicates that the Boschetti et al. (2019) and Padilla et al. (2014:2015) reference datasets were obtained from locations selected by stratified random sampling. For some of the other datasets, this is less clear. It would be useful for the authors to check each dataset and be sure that it is indicated whether the dataset had an underlying randomized sampling design.

2. Related to the previous comment, the manuscript identifies that several of the datasets included were selected by stratified sampling designs, and these designs had intensified sampling in high burned area strata. According to the original articles associated with these datasets, rather complex estimation formulas have to be applied to such data (i.e., the less complicated formulas of simple random sampling are not appropriate when the sampling was stratified with different sampling intensities in the strata). It would therefore seem necessary that users of these reference datasets be cautioned about the need to use proper estimation formulas if users are to correctly report accuracy from these stratified sample datasets. This would also create the need to include in the datasets the information required to apply these estimation formulas,

C2

for example stratum sizes, the stratum ID of each sampled unit, and perhaps additional information depending on the specific details of the particular dataset.

#### Technical Corrections and Suggestions

Throughout, readability would be enhanced by using paragraph indents at the start of each paragraph.

L23: insert “a” to revise to “requires a high level”

L26, L29: Given that the acronym BARD was defined at Line 26, replace “The Database” with “BARD”

L40: “sensors” should be “sensor”

L41: revise to “reference data that are based on” [“data” is plural so “data that are”]

L46: “products” instead of “product efforts”

L63, L79, L105, L106, L159, L161, L164, L198, L205, L207, L209: Throughout the manuscript, the words “file” and “files” are sometimes used to refer to the actual reference data. For example at L63, the “files” were not derived from pairs of images, but rather the “reference data” that are stored in the files have been produced from the pairs of images. The text should be revised to replace “files” with “reference data” unless the text is referring to the actual files that store the reference data.

L64: Replace “without probabilistic meaning” by “that were not selected using a probability sampling design”. It is not clear what “direct sampling” is. Is direct sampling convenience, purposeful, or other sampling without randomization?

For all examples at Lines 65-70, it appears that there was a rationale for why sites were selected (even if they were not selected by a randomized protocol). It would be useful to mention what purposeful selection criteria were used. The Roy and Boschetti example mentions sites selected to be spatially distributed across the landscape, so this is an example where the manuscript provides useful additional information regarding the

C3

purposeful selection criteria.

L70-71: If Boschetti et al. (2019) collected data for only a single year, does that qualify as a “full spatio-temporal validation”? It would be helpful to define what a “full validation” is in regard to time and space.

L88: insert “design” after “random sampling” to create “stratified random sampling design”

L89: Consider revising to: “Boschetti et al. (2016) extended the sampling design to include the temporal dimension of the sampling units.”

L90: insert “the” between “allocate sample” and delete “a” from “example a stratified”

L91: insert “the” before “sample”

L94: replace “are” by “is” because “dimension” is a singular noun

L99: delete “a”

L106: Consider revising to: “The procedures implemented to obtain those burn patches are diverse, depending . . .”

L109-110: Consider revising to: “Parts of the scene that cannot be observed or interpreted because of clouds or sensor problems (i.e., Scan Line . . .”

L115: replace “such” by “each” and replace “like” by “such as”

L153: Are n=127 and n=131 the number of TSAs sampled? It is not clear what these numbers represent.

L170: delete “to each sample unit” because this threshold is applied to all TSAs. That is, all TSAs are assigned to strata as part of the sample selection process. It is not just the sampled units that are assigned to strata.

L172: given that “proportional allocation” for stratified sampling is defined as the sample size in each stratum being proportional to the number of units in the entire study

C4

region belonging to that stratum, replace “applying a proportional allocation” by “applying a sample allocation”.

L182: replace “in” with “of” and replace “days” with “day”

L185: It is not clear how the actual time period covered by these “long units” is defined.

L186: Consider revising to: “Reference maps using long units concatenate consecutive 8-16 day maps (Fig. 5).”

L188: The 50 units are for fire CCI Africa compared to 100 units per year for FireCCI global?

L189: replace “consists on” with “consists of” and replace “perimeters” by “perimeter”

L190: replace “units” by “unit” (2 cases) and “days” by “day”

L198: remove “A” before “systematic sampling”

L201: replace “the whole” with “all” and replace “was” with “were”

L203: “consecutively” should be “sequentially”

L209: “joined” should be “joint” and “by” should be “between”

L213: delete “the” before “77%”

L219: replace “scar samples” by “scars sampled”

L223: “days” should be “day”

L224: “pair” should be “pairs”

L228-229: Continue to use the same phrasing as at L180 and L207 to identify the stage of the reference dataset. The sentence structure at L180 and L207 is much easier to read.

L231-232: replace “generated to perform the validation of the BAECV” with “generate

C5

to validate the BAECV”

L232: Move the text “Landsat Burned Area Essential Climate Variable” to before the first use of BAECV at Line 231.

L238: delete “A” before “systematic”

L239: the three values of n sum to 335 images not 336

L243: replace “. . . only two (pre and post-fire image. . .)” by “. . . only two images (pre and post-fire). . .”

L266-267: Continue to use the same phrasing as at L180 and L207 to identify the stage of the reference dataset.

L272: “wildfires” should be “wildfire”

L279: “were” should be “was”

L283: “postfire” should be “post-fire”

L284: “formers” should be “former”

L290-291: Continue to use the same phrasing as at L180 and L207 to identify the stage of the reference dataset.

L306: Consider changing “futures updates come to replace the lack. . .” with “future updates remedy the lack. . .”

---

Interactive comment on Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2020-74>, 2020.

C6