**Reply to short comment 3**

*Given the effort and cost involved in generating validation dataset, the standardization and documentation of existing datasets for future use is certainly a meritorious effort, and there is no doubt that the datasets made available by the authors of this paper will find a use in the fire community.*
*I have however some concerns.*
*1) Scope of the paper and qualification of the dataset The way the dataset is presented could lead to some confusion and misinterpretation. The paper title refers to the 'development of a standard database' and the abstract refers to the present work as the compilation of 'the first Burned Area Reference Database'. This is misleading, because the work described in the present paper is limited to the collation of existing datasets, through standard GIS operations described in Section 2.3, namely the conversion of the various datasets to the same file format (shapefile), the use of standardized file names and the creation of simple metadata (Table 1). The wording of the abstract, i.e. referring to BARD as 'the first Burned Area Reference Database', is incorrect, as this is not the first burned area reference dataset - all the datasets that constitute the BARD are pre-existing. Maybe 'the first publicly available burned area reference dataset' would be a more appropriate statement.*

Response:
We have followed your suggestion and changed the sentence to "the first publicly available burned area reference dataset". Actually, identical sentence was already included in the conclusion section: '*the first publicly available burned area reference dataset'* where we clearly mentioned that '*BARD is the first publicly available database that compiles and standardizes previously generated validation reference data*.'

*2) Degree of novelty Section 2.3 is the only section that reports original work (i.e. the conversion of all data to shapefile, the standard filenames and the metadata), while the rest of the methods document what was done by the various research teams in the projects that provided the data.*

Response*:*
The novelty of this paper is the compilation, standardization and public release of existing BA validation datasets, never done before.

*3) Sampling. Section 2.1 ('Selection of the validation sites') describes a procedure for stratified random sampling of the burned area reference data that was followed by some of the collated datasets (but not all), which is extremely confusing. An inattentive reader might be led to believe that the BARD dataset itself is the result of a stratified random sampling, rather than the collation of datasets some of which were the results of stratified random sampling (albeit with different methods) and some that are not.*

Response:
We have clarified this point in section 2.1 to avoid confusion. Please, note that table 2 summarizes the sampling methodology applied to each dataset.

*4) Stage 3 validation data set. Much is said throughout the paper of the compliance with CEOS Stage 3 validation requirements, but the BARD dataset as currently defined does not meet those requirements, i.e. it would not allow for the use of unbiased estimators of the accuracy metrics, and their associated standard errors. In the current form, pieces of BARD could be used for a Stage 3 validation, whereas other pieces could only be used for a Stage 1 or Stage 2 validation. Could the authors provide a harmonized statistical framework for the estimation of accuracy metrics from the whole BARD dataset?*

Response*:*
BARD is a compilation of datasets that have been produced in different projects where different methods were applied. Even those datasets produced in the throughout the life of the FireCCI

project present substantial differences. Please note that FireCCI project is a long term project that started in the year 2010 and, through the years, methods have been improving. That said, the aim of BARD is not to provide a harmonized statistical framework for all contributing datasets, because BARD is not a dataset itself but a compilation of datasets produced by different international projects and years. If that was the aim of BARD, we would only have made available the FireCCI global (2003-2014) dataset which is the one that covers the longest period but, instead, we choose to make all possible datasets available, and leave users the freedom to use the dataset or datasets that best suits their needs.