

## ***Interactive comment on “A database of global reference sites to support validation of satellite surface albedo datasets (SAVS 1.0)” by Alexander Loew et al.***

**Anonymous Referee #2**

Received and published: 12 June 2016

General comments: The article "A database of global reference sites to support validation of satellite surface albedo data sets (SAVS 1.0)" by Loew et al., describes a new tool to identify surface field locations that may be best suited for satellite derived albedo product inter-comparisons and evaluations. The database provides information on homogeneity of land surface type and an estimation of the variation of minimum/maximum annual NDVI (both relying on the 2010 300m ESA CCI land cover), and an estimation of the surface topographic variation all within 1, 2, 5, 10, and 20km of the site center. This database was specifically developed for evaluation of geostationary data and therefore is relatively lacking in high latitude sites. Overall, however, this data set represents a highly useful contribution to the field. The description is appropriate with only three

C1

major caveats which need to be addressed.

Specific comments: Firstly, the data set presumes that homogeneity of land cover (within 1-20km) is the most important attribute in satellite derived albedo product evaluation and inter-comparison. While this is important and interesting information, evaluation can not only be restricted to homogeneous regions (so-called pure pixels) but must also be applied to heterogeneous regions (mixed pixels) to fully characterize a satellite product. Therefore it is important to not only identify field sites situated within homogeneous land covers but also should identify sites which lie within heterogeneous land cover combinations that are similarly heterogeneous within the 1-20km range. One immediately thinks of mixed forests or savanna locations that would represent a mix of classes in the close vicinity of a tower with high quality in situ measurements but which might or might not be very representative of the same mixed forest over the greater region due to species mix, structural variability, canopy cover, and timbering. One must also caution that heterogeneous covers like mixed forests can be quite uniformly homogeneous over large areas during the growing season but can become quite heterogeneous during the shoulder and winter seasons – yet the site may still (or may not) be quite representative of that heterogeneity present during these seasons. This difference between site homogeneity and site representativeness needs to be much better discussed. Secondly, there is little assistance given to the user to reach the in situ data associated with each site other than the list of links in Table 2 of this document – more information should be including in the site information. This is particularly egregious with respect to the Baseline Surface Radiation Network (BSRN) sites which are explicitly identified in GTOS and GCOS documentation as the gold standard of in situ (tower) data for evaluation of satellite data products – but these sites (assuming they are representative) are not flagged as achieving GCOS criteria and standards in this database (note: Albedo and Reflectance Anisotropy, ECV-T8: GTOS Assessment of the status of the development of standards for the Terrestrial Essential Climate Variables, 2009 and GCOS document 154, pages 42 and 81). Furthermore there is little discussion of the in situ data associated with each site anywhere (other than the

C2

links in Table 2). The Ft Peck site is illustrative of this. It is listed as a member of the SURFRAD network. Nowhere is it mentioned that it is therefore also a member of the BSRN network (as SURFRAD sites are the US contribution to BSRN) and as a BSRN site, this indicates that in situ instrumentation are carefully and frequently maintained and calibrated and the in situ data are of the highest quality. Is this a case of a “duplicate” site which is mentioned in the text?? In the case of duplicates, all associated networks should be listed in the dataset descriptions – not just one network. Now, Ft Peck not only provides high quality data of a relatively homogeneous site but the FOV of the in situ instrumentation is also spatially representative of the greater area. Tower based albedo data (from a 10 meter tower) represents a field of view of ~127 meters on the ground. Using the same Román et al., 2009 indices on higher resolution data (e.g. Landsat), the Ft Peck in situ data has been found to be spatially representative of the surrounding 2km and therefore one can safely perform a point to pixel comparisons with medium resolution satellite data products. On the other hand, the Bondville site is also a SURFRAD site (and thus a BSRN site). However it is only listed as an Aeronet site in the database (note that by not listing the other networks you may be implying that there is no in situ tower based albedo data available). It is situated within a large region of cropland. However, despite the homogeneity of the cropland land cover type surrounding the site, it is perhaps one of the least spatially representative sites for satellite evaluation in the BSRN system. This is because the tower is placed on a grassy area in the middle of the croplands and thus the types of crops surrounding the tower frequently change. Furthermore the greater region is sporadically intersected by roads and drainage ditches, thus greatly challenging the ability of the tower data to capture the spatial representativeness of the region. Thus, while this site is a long term site with high quality instrumentation within a large area of homogeneous cropland, it is not a particularly good site to use for satellite-derived intercomparisons and evaluations. Now while acknowledging that the SAVS database has not undertaken to directly provide in situ data, some links to ALL the data networks associated with a site would seem warranted in each site description and certainly some further discussion

C3

(and references) on the caveats in using the associated in situ data should be provided to the user. Thirdly, the dataset relies on the 2010 ESA CCI land cover for homogeneity measures. This land cover is becoming dated and thus the land covers surrounding a site may well be changing. A GoogleEarth cutout is also provided for each site to aid in visual inspection of the site. A signal of historical snow fall potential and of fire potential is also provided –but there are of course many additional types of disturbance which may have occurred and things may have changed quite a lot since the 2010 CCI (and there should be some mention of the overall quality associated with the CCI). It is not clear how frequently the information associated with the sites in this database will be updated and some discussion on this point should be provided.

Technical corrections:

Page 1, line 25: Why are only geostationary and AVHRR albedo datasets mentioned and not the 16 years of MODIS and MISR products????

Page 2, Line 10: In situ data representativeness is briefly mentioned here but perhaps a better discussion of site data representativeness - as opposed to only site land cover homogeneity - is warranted. This is also where the BSRN network, with its extremely high quality of in situ measurements, should be acknowledged (as well references to both GTOS (GTOS document ECV-T8) and GCOS documentation (GCOS document 154)).

Page 4, Line 17-23 Awkward phrasing. . . perhaps: “The evaluation of surface albedo data products typically requires that diurnal variations in surface reflectances be taken into account. . .” and “Stringent requirements on the characteristics of a reference site to determine whether it is suitable for coarse scale surface albedo are therefore required:”

Page 6, line 25 – Duplicates should have all of their associated networks identified.

Page 7, Line 14 - A few words are required here – indicating how important it is to verify that you are not looking at snow (or a recent fire scar) before using a particular site at

C4

a particular time.

Page 10, sections 4.5 and 4.6 - caution readers that all evaluations are based on the aging and static CCI

Page 11, Line 30 – Not all required data information is currently provided – again the duplicate sites should have ALL of their associated networks identified so readers can locate the appropriate in situ data.

There are some minor awkward English phrasings throughout that should be addressed by a copy editor. . . .

---

Interactive comment on Earth Syst. Sci. Data Discuss., doi:10.5194/essd-2016-11, 2016.