Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2019-200-RC1, 2020

Open Access Earth System Science Data Discussions

Interactive comment on "Development of a global 30-m impervious surface map using multi-source and multi-temporal remote sensing datasets with the Google Earth Engine platform" *by* Xiao Zhang et al.

Anonymous Referee #1

Received and published: 3 March 2020

This manuscript introduced a global scale impervious surface map generated with multi-source remote sensing datasets, and comparative analysis suggested that the developed map outperformed the state-of-art land cover products. Despite producing a global impervious surface map using manifold datasets is an important contribution to the global land cover dataset, a major revision suggestion may be given from my side.

Major revision points are as follow:

Printer-friendly version

Discussion paper

1. The review of impervious surface datasets should be further improved. Here I recommend a few examples: Global Man-made Impervious Surface (GMIS) Dataset, Copernicus land monitoring surface – high resolution layer imperviousness (although this dataset only covers Europe continent, it can be used as training and validation sample source), NLCD imperviousness products, Global Human Built-up And Settlement Extent (HBASE) Dataset.

2. The scientific importance of your dataset should be enhanced. Demonstration of multiple dataset contributions to land cover classification was not straightforward. Data (spectral or radar) characteristics on different land covers (e.g. vegetation, impervious surface and bare soil) should be revealed in detail.

3. The accuracy assessment experiment should be improved and expanded to multiple urban landscape types (e.g. globally selecting validation sites in more bare soil prevalent cities and vegetation prevalent cities), so that readers can clearly understand how multiple datasets work in land cover mapping under varying landscape conditions (the same reason as the 2nd comment).

4. The training sample source/method may not be scientifically sound. GlobeLand30 was adopted as impervious surface training sample source, however, this global land cover product provides users with artificial layer but not impervious surface layer. The "impervious surface land cover" used in this study is actually a mixture land cover of vegetation, impervious surface and bare soil in urban area.

5. More explanations may be required to arguments in this paper.

Here I am left with a limited number of questions about the specifics of implementation and implications of the method and results as well as clarity of this manuscript, which I note below.

[Questions for "introduction"]

"However, Gao et al. (2012) explained that these coarse-resolution global impervious

surface maps were not suitable for many applications and policy makers at local or regional scales (Line 35)." This part is not understandable, why the previous impervious surface maps are not suitable for certain applications and policy makers? Could you please explain it with more straightforward instances?

"(Chen et al., 2015; Gao et al., 2012; Goldblatt et al., 2018; Gong et al., 2019; Gong et al., 2013; Homer et al., 2015; Li et al., 2018; Liu et al., 2018; Sun et al., 2017). Line 50". Because of their similar works as you did in this paper (i.e. global or regional impervious surface maps), it is necessary to give more introduction to previous land cover datasets, and to present the importance of your work.

"However, these land-cover products focus on the overall accuracy of the mapping of all land-cover types rather than that of impervious surfaces alone (Line 55)". It may be confusing here. It implies the land cover products that focus the overall accuracy deliver low quality impervious surface map. What difference existing between "focusing on overall accuracy" and "focusing on impervious surface alone"?

"However, the NUACI product had a relatively poor performance in terms of producer's accuracy (0.50–0.60) and user's accuracy (0.49-0.61). Therefore, an accurate impervious surface map at fine spatial resolution is still urgently needed (Line 55)". Why did you only mention accuracy of NUACI here? How about other land cover datasets?

"However, these spectral mixture methods can produce underestimates in areas whether the density of impervious surfaces is high and overestimates in areas of low density (Sun et al., 2017; Weng, 2012) (Line70)". Spectral unmixing technique may have underestimate and overestimate issues, but how about its overall or average accuracy when comparing it with pixel-level mapping approaches?

[Technical questions for land cover mapping process]

The "data preprocessing" and "mapping approach" was mixed up, which makes readers difficult to capture the point of datasets and classification methods, so I may suggest

splitting them into different sections.

Questions on remote sensing datasets for classification (Lines 120). Descriptions of purpose and necessity for different remote sensing datasets were not clear, better wording in “Datasets” may be required.

Do five data sources contribute equally to classification? How do they theoretically work for differentiating different land covers? (Line 120)

How does C-band SAR imagery contribute to differentiate impervious and pervious surfaces? How do artificial buildings, forests, grassland, and bare soil respond to SAR imagery? Please clarify it. (Line 130)

How do EVI imagery work for classification procedure? (Line 145) Why do you involve DEM dataset in land cover mapping? How does this dataset work in differentiating impervious and pervious surfaces? (Line 150)

Questions on introducing state-of-art global impervious surface products. (Lines 205). GlobeLand30 actually does not provide impervious surface land cover, please adopt other global land cover dataset instead (Line 160). Furthermore, the review of the published global impervious surface datasets should be improved. For instance, three important global (or continental) impervious surface datasets – global man-made impervious surface (GMIS) dataset, NLCD impervious surface layer and global human built-up and settlement extent (HBASE) dataset- were not introduced.

Questions on selecting training samples (Lines 205). The impervious surface training samples were selected based on GlobeLand30 map. However, GlobeLand30 only provides “artificial surface” which consists of impervious surfaces and small patch vegetation areas in urban area. Thus, the training samples of MSMT_RF could be no longer reliable although extra datasets were used for samples filtering.

The training samples for classifier may be collected from other impervious surface datasets instead of GlobeLand30 map.

It is not clear that how and why twelve sampling sites (i.e. high-density sites, medium-density sites and low-density regions) were selected? How do spectral features vary among these sites? What features were exhibited by different density regions? This information should be updated.

Moreover, data preprocessing procedures and mapping methods are mixed up, which makes manuscript confusing. I suggest separating them into two sections. In particular, more explanations of the preprocessing operations should be made,instead of only citing reference literature.

Some parts of introduction to data preprocessing were also not understandable. Here is an example: “the suburban areas or rural villages were also easy to confused with croplands (Li et al., 2015)”. It is not reasonable to compare a land use element (suburban area) with a land cover element (cropland). Explanations are always required for each of your arguments.

Questions on accuracy assessment (Lines 315). Two accuracy assessment was conducted respectively in “fraction” way and “classified pixel” way. How much difference do the two accuracy assessment methods make? What special information can be provided by each method?

Questions on Figure 5. As mentioned in previous questions, further review of currently available global impervious surface maps is needed. In the revised manuscript, I suggest adding error bars for progressive fraction intervals (e.g. 0.05, 0.1, 0.15, 0.2, . . ., 1.0).

Confusing parts in Section 4.2. “As the stratified random sampling strategy was applied to each validation region independently, the low and medium density regions were easier to select these mixed impervious validation points (simultaneously containing the impervious and non-impervious surfaces in the 30-m $\times$30-m validation window and the impervious areas exceed the predefined threshold of 50%) which were most difficult to identify for impervious surface mapping (Line 380).” What information did you want

to present?

Questions for Figure 6. It is clear to show difference between impervious surface maps but not clear to visually compare RGB pixels with your maps. The RGB satellite images of macro areas may not be suitable to compare it with classified land cover map. Subset urban areas are preferred so that readers can clearly see how well the map is classified. Besides, I may not agree that “low, medium, high- density” areas are representative for comparison. To improve the figure, I suggest globally selecting urban areas with different landscapes (e.g. desert landscape urban areas such as Phoenix city, vegetation prevalent cities such as New York City). Futhermore, please do more works in reviewing global impervious surface datasets.

[Questions for “Discussion”]

Figure 7 is an experiment result, and it should be moved to “Result”.

“The importance of all 37 training features for the six regions is illustrated in Fig. 7. These results indicate that the Sentinel-1 440 SAR features (VV and VH) had the greatest contribution to the final decision in most regions because SAR images can provide information about the structure and dielectric properties of the surface materials (Line 440)”. What VV and VH feature difference is revealed between different land covers (e.g. impervious surface, forest, croplands, bare soil, water)?

“Similarly, Zhu et al. (2012) demonstrated that the inclusion of multi-temporal imagery increased the accuracy by 8.9%. Schug et al. (2018) also found that bi-seasonal information could produce a more reliable performance than a single-year composited image. Therefore, temporal variability can be considered an important addition to accurate impervious surface mapping (Line 455).” Discussion and explanation should be made. Please exactly explain the theory in which how these datasets work for improving classification accuracies. Which land cover accuracy is improved by including these datasets?

“Similarly, Clarke et al. (1997) explained that topographical variables (slope, aspect and DEM) contribute a lot to impervious surface mapping. These features are, therefore, indispensable in the accurate mapping of impervious surfaces in complex landscapes (Line 465).” Clarke et al. (1997) was cited without further explanation. Readers would like to know mechanism of topographical variable contributing to impervious surface mapping?

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2019-200, 2020.