

Review ESSD-2020-157, terraces

Overall, very useful, a good product for ESSD. With some small efforts authors could improve data access and data description for many users.

Section 2.1.1: Authors downloaded a LandSat product that already included cloud masking. But they applied their own additional cloud masking? Please clarify. Why and how did they perform the second cloud masking.

Section 2.1.2: STRM DEM - provide a citation for the (standard?) void-filling process, here and/or in Table S2. Many readers will not know these acronyms?

Line 112, 113 “we suppose the terrain changes little in decades.” Interesting, this reviewer tends to agree. But what decadal changes might impact terraces? Consolidation? Urbanization? One supposes most terraces lie above floodplains, but construction of dams for agricultural water or flood management? In a dynamic Chinese economy, can we really assume static distribution and abundance of terraced agriculture? With their expertise, perhaps authors can and should comment? I note this as possible uncertainty because a few paragraphs later the authors mention use of / calculation of Normalized Difference Building Index (NDBI), which implies that they might need to account for temporal changes in land use over their time periods?

Section 2.1.3, GlobeLand30 represents an interesting product, perhaps unfamiliar to global LULUC communities. Need an actual citation? GlobeLand30 is not in GEE? Same issue about static vs dynamic: in rapidly-changing Chinese economy, can one accept the assumption of no change between 2010 and 2018?

The reviewer confirms 39 features from Table 1: 7 spectral bands from LandSat plus four indices, each with three ranges, plus 6 elevation-related features from SRTM. But several indices listed, e.g. NDBI, incorporate middle infrared (MIR) while MIR not listed among the bands downloaded or processed. Please can authors explain?

Line 153, 154: “(2151 terrace samples and 2639 non-terrace samples) were collected by visual interpretation of Landsat images, SRTM DEM data, cropland extent data extracted from GlobeLand30, and Google Earth images”. Authors inspected nearly 5000 images visually? To confirm terrace vs non-terrace? To confirm other land use (agriculture) feature? Impressive but difficult to understand or replicate? What did inspectors gain from GE images not available from LandSat or SRTM? Google Earth images within GEE or separate? GE images one of the least replicatable aspects of this research?

Lines 162, 163: test dataset has a much higher ratio on non-terrace to terrace (almost 10 to 1) than training data (more like 5 to 4). Does this matter? Will this difference arise later as an uncertainty factor?

Section 2.4 Terrace/non-terrace. Because not more than 20% Chinese land area = agriculture, terrace area also has to be << 20%? RF classifier specific to each province - very positive! All this done within GEE?

Section 2.5 Post-processing. Understand the purpose for and mechanism of mode filtering. More interested - and worried - about the sieving process. Terraces on slopes often have a long axis following contour lines but a short axis up- or down-slope. Which axis and/or what areal dimensions do the authors use here as “small”. “Small” area for terraces not the same as “small” area of non-terraces? Confused about pixels vs resolution. Understand LandSat at 30

m, but that means 10 pixels represents 300m? Rather large dimension for a “small” area? Or perhaps not on 1km x 1km scale? Please correct my mis-impressions.

Line 204: Because, in subsequent text, authors often switch between user/producer terms (UA, PA) and commission/omission terms, take this opportunity to define both? For example, in line 204, “the user’s accuracy (UA) of the terrace class” could become ‘the user’s accuracy (UA, also referred to as ‘commission error’) of the terrace class’. Authors will know how best to make these connections but because they use both (user/producer and commission/omission) they need to clarify.

Line 211 - again, reader needs to know explicitly how authors define “small”.

Line 212 - is this second test set of 301 samples also available at Zenodo? Not obvious? It seems like the authors regard it as an important independent test product?

Line 240 - if cropland represents 20% of total China land area, and terraces represent 26 to 28% of cropland, then terraces represent 5% to almost 6% of total China area? I do this as a sort of ‘mass balance’ check; have I got this correct? If not correct, what did I miss?

Line 277 - finally a list of 12 excluded provinces! Having this information appear earlier in the manuscript would have answered many questions for this reader.

Uncertainty assessment/discussion and feature importance, including Figures 6 to 10, represent strong positive features of this data description. Thank authors for this careful analysis.

Limitations and directions - overall a very useful discussion, particularly about spatial resolution. Again this issue of visual inspection. Who has energy to visually inspect 1000s of images and how does one quality-control outcomes of such inspection? Even if the authors can not provide quantitative assessment, they could help other users with a general estimate of the effort involved? The authors also discuss merits of GlobeLand30 vs other LU products. Again, users will want to know how they might get access to GL30 (e.g. via GEE?) as well as other references to GL30 use and accuracy.

On Fig 3, not obvious that Xinjiang or Heilongjiang provinces have terraces. Need different or brighter color scale. Could authors mark location of positive outcomes shown in Figure 4 in Yunnan, Hunan or Guangxi provinces on Figure 3?

Figure 5 does not include all the provinces show in map on Figure 3? By cropland or terrace abundance or some other LU factor, authors have eliminated far west or far north provinces from their analysis. Did the reviewer miss an earlier statement to this effect? These represent the “other provinces” mentioned in line 238? No, they represent the 12 province excluded, listed at line 277. A reader needs to see this exclusion information earlier, before any results, perhaps even before most methods?

Figure 10: very difficult to read / interpret axis labels. Table S1 helps, it should move into main manuscript as part of legend for Figure 10?

(Because Table S2, the source attribution table, also belongs in main manuscript, perhaps in or near Section 2, authors could include S1 and S2 in main manuscript and thereby eliminate supplement?)