Review ESSD 2018-14, high res imagery of Lake Tana basin

Many errors in text, authors need systematic reading and correction.

Page 5 line 113 “aerial” Over most of the land masses of the planet Google Earth does NOT ingest aerial (e.g. airborne) but only satellite sources. If the authors know of specific aerial images included in specific Google Earth products, they should identify those sources and those images. Otherwise this statement seems inaccurate.

Page 5 line 118, classification system. Why did the authors not use IGBP land use categories? Their list here does not match other classification systems and therefore does not allow external comparisons. If they have valid scientific reason to keep these categories, they should show how their categories cross-match to IGBP categories.

One assumes - for lack of detailed information from the manuscript - that these authors have used very high resolution Google Earth images, e. g. at or better than 5 m resolution. But: a) most Google Earth images do not include proper metadata, e.g. source files (LandSat, SPOT, etc.), processing techniques, exact times, actual resolution, etc.; and b) recently Google has started to mix in older LandSat images - of lower resolution - in their time sequences at many locations. These authors do not assure us which images they used and at which resolution. They should look at Pinzon and Tucker 2014 (Remote Sensing 2014, 6(8), 6929-6960; doi:10.3390/rs6086929), and particularly at panel E of Figure 4 in that manuscript. What, if anything, have they provided here that improves on those prior AVHRR-based NDVI data? I agree that they might have a better product but they have failed to demonstrate how or on what basis.

The manuscript lacks a serious discussion of methods, validation and uncertainties. It does not provide guidance or assurance for other users. The entire Results section relies mostly on previous surveys with little reference to this work. The authors have largely avoided and ignored the largest land category, cultivated lands.

In addition to identification errors as pointed out by the other reviewer, this reader expects two other sources of error: changes in the satellite sources applied image by image by Google Earth (which the authors have not mentioned and perhaps can not in fact extract from the GE images themselves); and survey / validation uncertainty. Other on-the-ground land use surveys involving human observers (e.g. see the Woods et al. survey of landscape features in the UK, https://doi.org/10.5194/essd-10-899-2018) describe in great detail observer training, paper-based and digital support tools, digitising and annotation errors and exceptions, subsequent replication by experts, etc. We get no idea from this manuscript how or based on what criteria they involved their expert(s). If the authors do not understand and address uncertainties, readers / users get no basis on which to build trust in these polygons and categories.

Most of the results section seems drawn from prior literature rather than from analysis of these data. Given errors, uncertainties, and the decision to avoid cultivated lands, this reader doubts that the authors can justify and document discussion in four significant figures (e.g. natural forest at 0.82% or plantation forest at 1.92%). In places these authors invoke temporal changes in land use but they provide absolutely no data here to document those changes.

The data downloads cleanly from 4TU but the .txt file only recites the categories without providing additional useful information or metadata, at least some of the files for the individual categories open only in ESRI ArcGIS, and users get no access to summary information they would need to check the authors’ results.

Given substantial absences of information in this manuscript and data as presented, this reviewer does not agree that (line 292) “this vegetation map could offer reliable information for vegetation conservation in Lake Tana basin”.