Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2017-22-AC3, 2017

Open Access Earth System Science Data Discussions

Interactive comment on "Spatial and temporal patterns of plantation forests in the United States since the 1930s: An annual and gridded data set for regional Earth system modeling" *by* Guangsheng Chen et al.

Guangsheng Chen et al.

chengu1@auburn.edu

Received and published: 29 June 2017

Overall Quality: The goal of developing a spatial dataset of historical forest plantation activity is worthy and could lead to valuable data for a wide array of scientific studies. However, this manuscript is difficult to understand and perhaps is fundamentally flawed in approach. Of primary concern is incorrect assumptions embedded in the computational technique (Equation 1; see specific comments below). Beyond that, the approach results in a dataset of unknown quality and misleading in what it represents (basically backcasting the spatial distribution of historical plantation activity based on a

spatial distribution from a single point in time). Furthermore, the approach used to aggregate fine-scale data to a coarse resolution (8 km grid cells) is not detailed, leaving, the user/reader uncertain as to what is represented in the final dataset.

Response: Thanks for the constructive comments and suggestions. We have substantially revised the manuscript. There are a lot of plot/stand-scale studies for the effects of plantation forests on carbon, nitrogen, water and greenhouse dynamics, but still lack of a scaling up from these field observations due to lack of a gridded and continuous plantation distribution data. Our purpose of this study is thus to develop an annual gridded forest plantation dataset, which will be readily applied in the regional/national scale ecosystem modeling to estimate the effects of forest plantation on terrestrial biogeochemical and hydrological cycles. This dataset does not target and also unsuitable for watershed/stand/plot level studies due to the lack of a high spatial accuracy and resolution. The reason with coarser resolution (8 km) is a combined consideration of the limitation of FIA plot numbers and the consistency of regional modeling environmental datasets, which are also explained below. The spatial resolution can be reduced to the actual plantation management size if more FIA plots for plantation distribution are available in the near future.

Our approaches and assumptions are commonly used in generation of long-term environmental data especially land use data for regional/national/global modeling studies. The generated data especially historical spatial patterns have many uncertainties, which are resulted from the limitation of available data inputs. The reason for the assumption is also because we only have a single data point for fractional plantation data. These inadequacies are uncertainties rather than method flaws or errors, and can be reduced in the future with technology improvement or more available data inputs. In spite of the uncertainties, our dataset will still be useful and can improve our understanding of biogeochemical and hydrological cycles of forest plantations. Our work is an initial attempt for providing the first gridded long-term plantation data.

We have more specifically addressed these questions point-by-point at below.

Specific Comments: The fundamental allocation of FIA plot data to grid cells appears to come from Equation 1 (line 149). The authors are slightly inaccurate in their description of fuzzing and swapping of plots in lines 140-146. FIA "fuzzes" the coordinates of plots by adding a random error to published coordinates. The "swapping" is a further step in which the measurements obtained at one plot are "swapped" with the location of another plot with similar ownership and geography. The 675 meters cited by the authors was the amount that plot coordinates were rounded to in the Eastwide Database (EWDB), the mechanism for FIA data delivery prior to around year 2000. Currently, the policy is that plots are "fuzzed" to within about 0.5 miles (800 m; see O'Connell et al., 2015). Regardless, the authors use the 675 m as the "represented area" of a plot. This is incorrect. FIA computes the area represented by a plot based on sampling design and forest area estimates; this area is usually about 5,000 to 6,000 acres (in the South). Thus, the actual area represented by an FIA plot (about 2,400 ha) is more than 50 times the value used by the authors in Equation 1 (675 m x 675m = 45.56 ha). Perhaps this error necessitated the adjustment mechanism used by the authors to force the gridded estimates based on this equation to match state or county estimates (section 2.7 of the manuscript).

Response: The fuzzy distance is about 0.5-1 mile (O'Connell et al. 2015), and there are about 20% of the plot locations are swapped due to privacy. In the revised version, we apply 1 km to represent the inventory area of each plot and generate a new plantation fraction map at 8 km resolution (see the new Figure 4). In fact, under the condition that the fuzzy and swapping distance is less than 8 km, the exact value of the fuzzy and swapping distance is not very important in this study. We don't intend to generate an accurate plantation fraction map (hard to generate since many plantation forests are not included in the FIA inventory); instead, this map is a middle level data used to identify which 8 km grid cells have more inventory plots of plantation forest, with more plots the higher probability to be identified as "plantation". The useful information is the relative magnitude, so our generated distribution pattern of plantation is the same under both conditions.

Lines 233-236: The authors chose to provide datasets at 1 km and 8 km resolution. Why were these chosen? Average pine plantation areas (stands) in the US South tend to be around 50 acres (20 ha). Thus, an 8 km grid cell (6,400 ha) could contain hundreds of planted stands, of varying ages and species. The 1 km resolution is much more reasonable, given the precision of the data used (e.g., 250m pixels for forest species). It is questionable whether land use dynamics that operate at a 20-ha scale are adequately represented in 6,400-ha data.

Response: 1 km resolution is used for the plantation fraction data since the representation of each FIA plot is around 1 km. We also thought about the average pine plantation size in the US, so we originally plan to produce a plantation distribution map at 1 km spatial resolution. But we have too few inventory plots (16,677) and thus we can not get sufficient plot information to identify the plantation area at the 1 km spatial resolution (total plantation area in 2012 is 268,279 km2; we are unable to identify all the plantation distribution grid cells with the limited plot numbers). Alternatively, 8 km spatial resolution perfectly match with the FIA plot numbers and representing area. In addition, we choose 8 km resolution (about 5 arc minutes) because this is a relatively high resolution for ecosystem models to simulate plantation impacts for the CONUS. Many previous modeling simulations used 0.5d (around 60 km) or 0.25d (around 30 km) to simulate the ecosystem responses to environmental changes in the CONUS. We also considered that the PRISM climate data is also 8 km and thus 8 km plantation data could be suitable and readily used by modelers in the future. We add a sentence in the method to explain the choice of 8 km resolution.

We understand that 8 km is too coarse as compared with the average plantation size (20 ha). Due to the difficulty mentioned above, we have to choose to lose some spatial accuracy. An 8 km grid cell may also cover several plantation management units. A common and needed assumption is generally used in generating long-term land use history data for model simulations to match the inventory data. That is, the fraction of this land use types in the neighboring grid cells are moved to and converged in this grid cell. For example, in this study, we assume that all the plantation area in the neighboring grid cells are concentrated in this identified grid cell. As such, the plantation area in this identified grid cell is used to represent the mean responses of surrounding grid cells. As we also mentioned in the manuscript, unlike other land use and cover categories (e.g., urban, forest, grass, shrub), all the present remote sensing images from different sensors are unable to accurately classify plantation forests. At present, the spatialization of plantation area has to heavily rely on the inventory data.

The approach described by the authors (unless I misunderstand) seems to use the gridded distribution of plantation area from 2000 to 2004 to compute percent of a grid cell composed of forest plantations. This same spatial distribution is used to apply to regional/state estimates of plantation area for all other years. Thus, while this approach is based on available data, the implication for users of this dataset is that the land use dynamics of 2000-2004 are the same that have existed since 1924. This is highly unlikely, and gives a false impression to users of this data. For example, pine plantations have existed in the coastal plain for decades, but generally have given way to encroachment of development in rapidly urbanizing coastal areas such as Charleston, SC and Jacksonville, FL. Similarly, there have been back-and-forth dynamics between planted forest and agriculture, driven by incentives to landowners and crop and timber prices. A spatial dataset of historical plantation locations based on current distributions will mislead users that don't take the time to understand the approach used to develop this data.

Response: We understand that the spatial distribution would be biased due to many factors such as cropland expansion/abandonment and urbanization. Multiple time periods of gridded plantation forest fraction data could be perfect to delineate the historical distribution patterns. Under the limitation with only one time-period fractional data, we try to reduce the spatial uncertainty through integrating the state-level plantation area, subregional annual planted tree area data, and forest distribution map. These auxiliary data can at least restrain the spatial biases within a state boundary. For example, the

plantation forests, which should be located at the coastal area in SC, could be shifted somewhere but still located in the SC. The assumption that plantation forests don't convert back to natural forests or other land use is also because we only have one period fraction data (if we have multiple time points, we don't need to have this assumption). This assumption is commonly used to generate long-term historical or future land use data. This is also the major uncertainty (not a flaw or error) in our data generation.

There is confusion in the paper between annual rates of plantation and cumulative area planted. Figure 5 and Figure 8 clearly show different data (Y-scale is very different), yet the figure captions both mention "annual" planted area.

Response: These two figures indicate different datasets. The Figure 5 shows the annual planted tree area, while Figure 8 shows the plantation distribution area. It is a little confused between "annual plantation forest area" and "annual planted forest area", with "annual plantation area" representing the status of total plantation area while "annual planted area" representing the annually newly planted tree area. To make it easier to distinguish the two concepts, we replaced all places of "annual planted forest area" with "annual planted tree area", and delete "annual" in the "annual plantation forest area". Hope this term can be less confused.

The Figure 8 Y-axis and captions are revised (see the new Figure 8). The caption is changed to "Total plantation forest area (1000 km2) for different regions in the CONUS during 1928-2012".

Clearly, plantations on a 20-35 year rotation will turn over multiple times during the 80 years of this dataset, making cumulative estimates meaningless.

Response: We delete the whole paragraph (Line 299-307 in the old version).

Finally, much of the discussion in Section 3 pertains to trends in planted area, species distributions, etc., that are inputs to this research, not results. Planting rates, spatial distribution patterns, and management practices have been widely discussed in the forestry literature; the authors should either briefly summarize and cite other primary literature or leave much of this discussion out.

Response: We delete some of the discussions and add citations for some analyses (see the change-tracking version). We combined several datasets to generate a continuous annual datasets for plantation area and tree species. In addition, some of the data in our results are slightly different from the inputs. So it is still necessary to keep most of the analyses of plantation area, distribution and change trend.

Technical Corrections: The manuscript will need considerable editing for language/syntax correction.

Response: We have made substantial revisions throughout the manuscript (see the change-tracking version).

105: Fig. 1 shows, not showed.

Response: It is revised.

Lines 117-123 are not needed as there is a figure showing which states belong to the various regions.

Response: These sentences are deleted and only mention the 8 subregions here.

Line 130: “despite of some inaccuracy”; strike “of”

Response: Thanks for pointing it out. “of” is removed

Line 149: what is the meaning of the subscripts mn?

Response: m and n are used to represent the grid cell id (latitude & longitude, respectively). Considering they are not necessary to show the id here, I delete those “m” and “n”

Line 170: Oswalt et al. include data for all regions for 2013; this line should state 2014-2011.

Response: Thanks for pointing it out. We have revised from 2003-2011 to 2004-2011.

Line 171: Figure 6 should be Figure 5.

Response: Thanks. We have revised it.

Line 171-173: The authors discuss increases in annual planting area. They should either describe the reasons (Soil Bank programs, policy incentives) that are well documented in the literature or omit this.

Response: We deleted this sentence; instead, we add a short discussion in Section 3.1 to address the effects of incentives on the change patterns of plantation forest area.

Line 203: Figure 7 should be Figure 6. Line 241: Section 2.3 should be 2.4 Line 242: Section 2.2 should be 2.3

Response: Thanks for pointing these out. All these places are revised. After the initial review by the subject editor, we removed a figure and added a paragraph. We missed to find out these places for revision and thus caused these mistakes.

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-22, 2017.