RC2

General comments

The article is composed overall well and makes a useful support for the publication of the dataset. The method of using higher resolution, specialized data sets is appropriately chosen to refine the sometimes very vague class definitions of the ESA CCI land cover time series. The additional extension of the CCI user tool in order to enable the user to translate the CCI land cover classes to individual PFT maps addresses the needs of the regional climate model community, where different model families have different requirements to the land cover input.

The significance of such a dataset is paramount for the climate modelling community. The integration of the information of multiple high-resolution, remotely sensed datasets into the well-known ESA-CCI land cover time series certainly increases the potential high quality of the PFT time series. **However, all additional input datasets as well as the baseline ESA CCI incorporate uncertainties which are partially mentioned in the original dataset publications or investigated and published by the user community and should be at least mentioned in the present work. Therefore, I would suggest focusing section 3 more on the dataset accuracy aspect then on the comparison to the original PFT_{global} distribution.**

It is found that the cross-walking uncertainty is higher than the land cover product uncertainty itself (Hartley et al. 2017). Yet what is missing is an investigation of the quality of the final product. In addition to the use of the newly developed PFT dataset into RCM experiments and the comparison to the original ESA PFT cross-walking results, a validation through comparison to independent data should be an essential part of this effort. For example, within the GLOBCOVER initiative, the product was compared to a dedicated reference database (Defourny et al. 2009).

Note that a quantitative validation of the CCI medium-resolution land cover class dataset is available (C3S PQAR for LC, full reference given below). In building the PFT dataset from ancillary datasets, we take advantage of the high quantified accuracy of the land cover dataset and align the PFT fractional covers with the expectations for each class according to the class legends. Below, we provide some additional flavour regarding how well the ancillary datasets align with the class legend & the frequency with which adjustment to the PFT percentage was necessary to achieve alignment. We select the tree cover classes for this analysis.

Using all 300 m pixels of the CCI MRLC dataset that fall within the extent covered by the 30 m Hansen et al. 2013 tree cover dataset, we compared (1) the initial tree cover percentage at 300 m estimated from the ancillary Hansen et al. dataset (that is, before applying the harmonization procedure that aligns the tree cover fraction with the CCI MRLC class legend) with (2) the expected tree cover percentage based on the class legend. We additionally calculated the mean tree cover percentage across all pixels of a class, based on the ancillary product. We performed this analysis by class for each of the tree cover classes & subclasses. Results are shown in the table below.

Considering the mean tree cover percentage for all pixels within a class, based on the ancillary Hansen et al. 2013 product, only class 82 has a mean tree cover (3%) that falls outside of what is expected based on the class legend (in this case, 15-40%); however, this class has less than two dozen pixels globally.

Class 50 has the largest number of pixels among any of the tree cover classes and exhibits especially strong correspondence between the class legend and the ancillary data product, with 97% of pixels of this class having a calculated tree cover percentage that meets the class legend expectations. The ancillary data product suggests a high mean tree cover percentage of 89% for pixels of this class. (Interestingly, this is quite close to the 90% tree cover suggested by the original global cross-walking table.)

Classes 60, 61, 70, 71, and 81 each have agreement between the legend and ancillary product for at least 80% of pixels. Classes 72 and 82 have low levels of agreement (9% and 10%, respectively), but each of these classes has only a very small number of pixels. Both are subclasses with an expectation of 15-40% tree cover, with the majority of pixels that exhibit a mis-match having a tree cover percentage from the ancillary product that is lower than expected by the legend; in such cases, the alignment process increases the percentage tree cover so that it falls within the range suggested by the legend (see manuscript for method). Class 62 likewise has a legend expectation of 15-40%, but the mis-matched pixels show a more even split between over- and underestimation from the ancillary product.

Class code	Class description	% of tree cover class pixels belonging to this class	Mean tree cover percent across all pixels of this class, based on ancillary product	% of pixels having tree cover estimated from ancillary product matching legend
50	Tree cover, BE >15%	23.1	89	97
60	Tree cover, BD >15%	13.7	61	86
61	Tree cover, BD >40%	1.9	58	80
62	Tree cover, BD 15-40%	6.6	29	51
70	Tree cover, NE >15%	19.0	59	85
71	Tree cover, NE >40%	7.8	65	87
72	Tree cover, NE 15-40%	<0.01	22	9
80	Tree cover, ND >15%	19.2	33	62

81	Tree cover, ND >40%	<0.01	73	83
82	Tree cover, ND 15-40%	0	3	10
90	Tree cover, mixed tree type	6.4	79	NA
160	Tree cover, flooded - fresh or brackish	1.9	70	NA
170	Tree cover, flooded - saline	0.4	52	NA

A full comparison of the PFT maps with external data is beyond the scope of this paper, but would be a worthy topic for a follow up paper.

C3S PQAR for LC: Copernicus Climate Datastore the Product Quality Assessment Report ICDR Land Cover 2016-2020. https://datastore.copernicus-climate.eu/documents/satellite-land-cover/D5.2.2_PQAR_ICDR_LC_v2.1.x_PRODUCTS_v1.0.pdf

The article presents the workflow with all necessary detail for the user community, which makes the article quite extensive. **For a better overview a graphic outline of the general workflow would be highly beneficial for the reader.**

We have added a new table (below) to the manuscript that serves as a diagram of the method. We introduce the table at L222: “Table 2 is a high-level overview of the method used to derive the PFT fractional composition for the static pixels.”

Table 2. Summary of method applied to derive pixel-level functional type composition by land cover class. See Table 1 for more comprehensive class descriptions. PEA16 = surface water data product of Pekel et al. 2016. HEA13 = tree canopy cover product of Hansen et al. 2013. PEA13 = Global Human Settlement Layer from Pesaresi et al. 2013. PEA21 = tree canopy height dataset of Potapov et al. 2021. For the calculation of tree percentage: “Method 1” indicates that, in cases of disagreement in tree cover percentage between the ancillary dataset and the class legend, a window of up to 0.5° x 0.5° is used to estimate the final tree cover percentage based on neighbourhood pixels of the same class; and “Method 3” indicates that an upper limit of 14 % tree cover is applied based on the class definition. See the text for additional details about the processing and use of the ancillary data products, the method used to align the derived PFT percentages with the class legend, the scaling method applied in cases where the sum of PFT percentages from the ancillary data exceeds 100% in a pixel, and the method used to derive the PFT fractional composition for pixels falling outside of the extents of the ancillary datasets.

<table>
<tr><th>Class description</th><th>Inland water %</th><th>Tree %</th><th>Tree type</th><th>Grass %</th><th>Grass type</th><th>Shrub %</th><th>Bare soil %</th><th>Built %</th><th>Snow/ice %</th></tr>
<tr><td>Rainfed cropland (10-12)</td><td rowspan="10">PEA16</td><td rowspan="6">HEA13</td><td rowspan="6">Neighbourhood majority</td><td rowspan="10">100% - water % - tree %</td><td rowspan="3">Managed</td><td rowspan="10">0%</td><td rowspan="10">0%</td><td rowspan="10">0%</td><td rowspan="10">0%</td></tr>
<tr><td>Irrigated or post-flooding cropland (20)</td></tr>
<tr><td>Mosaic of cropland and natural vegetation (30)</td></tr>
<tr><td>Mosaic of cropland and natural vegetation (40)</td><td>Managed & natural mixture</td></tr>
<tr><td>Mosaic of tree/shrub and herbaceous (100 & 110)</td><td rowspan="6">Natural</td></tr>
<tr><td>Grassland (130)</td></tr>
<tr><td>Broadleaved evergreen tree cover (50)</td><td rowspan="4">HEA13, Method 1</td><td rowspan="4">Class legend</td></tr>
<tr><td>Broadleaved deciduous tree cover (60-62)</td></tr>
<tr><td>Needleleaved evergreen tree cover (70-72)</td></tr>
<tr><td>Needleleaved deciduous tree cover (80-82)</td></tr>
</table>

Mixed leaf type tree cover (90)			Neighbourhood majority						
Flooded tree cover (160-170)									
Lichens and mosses (140)		0%	N/A	100% - water %					
Sparse vegetation (150-153)		HEA13, Method 2	Neighbourhood majority	Tree % + grass % must be in range 4-14%			100% - water % - tree % - grass %		
Shrubland (120-122)		PEA21	Biogeographical approach	100% - water % - tree % - shrub %		PEA21	0%		
Flooded shrub or herbaceous cover (180)									
Urban areas (190)		HEA13	Neighbourhood majority	100% - water % - tree % - built %		0%		PEA16	
Bare areas (200-202)				0%	N/A		100% - water % - tree %	0%	
Inland water bodies (210)				100% - water % -	Natural		0%		

				tree %					
Ocean (210)	100%	0%	N/A	0%	N/A				
Permanent snow and ice (220)	0%								100%

Specific comments

L197 Sections 2.1.7 and 2.1.8 are missing, please adjust section numbering

Thank you for spotting this. Section numbering was adjusted.

L250f (also L375f) please explain a bit the size of the 0.25° neighborhood window, would a rather smaller window not be more appropriate to the ~300m (and finer) dataset resolution? Did you test smaller sizes?

We selected a window size of 0.25° as an appropriate size for picking up average features of the land cover. We wanted to avoid using a window that was too small to avoid propagating non-representative features of the landscape.