

Earth Syst. Sci. Data Discuss., referee comment RC1
<https://doi.org/10.5194/essd-2021-251-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.



Comment on **essd-2021-251**

Anonymous Referee #1

Referee comment on "High-resolution land-use land-cover change data for regional climate modelling applications over Europe – Part 1: The plant functional type basemap for 2015" by Vanessa Reinhart et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-251-RC1>, 2021

The manuscript presents a dataset that makes an attempt to bridge the gap between land cover maps (from ESA CCI) and the requirement of users (i.e. Regional Climate Models or RCMs), namely maps of Plant Functional Types (PFTs). Arguably, the solution from ESA CCI LC to do so, based on the standard cross walking table (CWT) of Poulter et al. (2015), can be improved. The main idea of this study is that these CWT can be regionalized by climate zones (the Holdridge Life Zones or HLZ) to be better tailored to the needs of RCMs, at least over Europe. The authors further go to do a validation based on a new point based dataset (GT-SUR).

I think the overall rationale for the paper is there. However, I find several issues which would preclude publication in the present form. Hopefully the authors will be able to tackle them and present a revised version that addresses these issues.

One first issue is based on the style. The document does not read super well, partly because it is being very descriptive on a case by case basis with respect to the classes. The text is too long in this respect and many of the details are not so relevant. The results were each class in analyzed separately with respect to filtering (e.g. figures 6, 8, 10-14 and related text) are very tedious. I suggest this should be condensed in some form or another, if not removed and relegated to annexes. There are other issues too that make the reading difficult (not in terms of language, but style), some of which I detail in the remarks further below.

Another major issue I have is that the validation based on GT-SUR is done only on the LANDMATE PFT map, and not on the baseline ESA CCI LC map, nor on the ESA CCI LC map transformant to PFT with the standard CWT. This makes it impossible to judge on the improvement in quality provided by the present study. It may very well be that all the accuracy with respect to GT-SUR is from the correct classification of the LC in the first

place, and not at all by the added value provided by the study. I suppose this is not the case, but it is not proven by the study. In fact, readers might think the authors are misleading the audience by claiming accuracy when none is due. To defend themselves from this, I strongly suggest that the authors engage in making the same comparison with GT-SUR with some baseline reference before they apply their contribution (and I would suggest to do it both for the LC, and for the PFTs with the default CWT). In that way, they will be able to isolate and quantify the added-value of their contribution

Another point that I think needs improvement is the whole part on the 'filtering'. It is not fully clear to me what is being done or why. It needs to be explained better and probably simplified. Also, all this part (along with the results) occupy a very substantial part of the manuscript and it is not fully clear why. Also, why are things aggregated at 2.5 dd instead of done at the 0.1dd? I fully understand that there are mismatches between the samples at point level and grid level, but I am not convinced by what is done here. I probably misunderstood something, but again, it should be crystal clear why things are done

Last major point is that I do not see why the dataset is done only for 2015. Especially since the ESA CCI LC maps are done for 1992-2020 now, and that they had been designed to be consistent in time. I think most users would expect you to exploit this info and provide not just a single date basemap but also a set of consistent maps for the observation period.

The rest of my remarks are here below:

L7: "to achieve the *best* possible representation"? Arguing it is the best is perhaps exaggerating. Unless you prove that cannot be anything better.

L31: I think it may be misleading that the concept of PFTs is increasingly used. It has been used for a long while now, and it is unclear if there is a trend in using it more. Some models are perhaps drifting away from it, or at least evolving in a bit of a different way... see: <https://doi.org/10.1111/gcb.13910> and <https://doi.org/10.1029/2018MS001453>

L38: as it seems clear the CWPs will be a key concept in the rest of the paper, it would be good to describe here a bit more what it is. The more general reader would not necessarily understand why a specific "procedure" is needed for "translating" LULC to PFT. I would further recommend to spend more time explaining the fundamental difference between LULC and PFTs. Thirdly, I would go forward and cite studies that have actually employed a CWP in practice (<https://doi.org/10.1038/s41467-021-24551-5> and

<https://doi.org/10.1038/s41467-017-02810-8> come to mind, there must be others). A separate dedicated paragraph in the intro on all of this would be warranted.

L48: is it specific for the RCM community? Could it not be applied more generally to the ESM (Earth System Modelling) community, irrespective of whether it is region or global?

L75: While I understand the logic of having a general workflow sub-section first to present an overview, the way it is it actually complicates the understanding. For instance, you talk about the HLZ map (L84) without saying before what it is (or rather describing what it is). It feels that maybe this section is not so useful as it is... if you want to keep it, it should be made much more generic so that it does not depend on what you will explain after.

L87: Not clear here (even if it may become clearer after) why an individual CWT is needed for each LULC class. The idea of a CWT would (a priori) be that you have one table for all.

L117: This simplified downscaling by bilinear interpolation will probably lead to problems in mountain areas. Any way to mitigate this? Could you comment on the possible consequences?

L127: typo on "merged"

Figure 2: There a logical inconsistency in the colour gradients you use. To be consistent, please always keep "deserts" as either the lighter or darker colour in each scale

Table 3 is a bit pointless as a table, as it only has one field (the number is not very meaningful in the context of this paper) and thus is just a list. If you keep this as a table, could you please add more fields with info (e.g. description)

L149: It is not clear enough where the information to make the C3/C4 distinction comes from. This separation is very important for modellers, yet the CCI LC product (and standard CWT) does not provide any way to do the distinction. This suggests it is difficult to do. Claiming that here you can do it would require more details on how it is done and some validation that is it done correctly. Furthermore, could you also clarify in what ways this is better than using the Still maps commonly used by modelers (e.g. <https://doi.org/10.1029/2001GB001807>)

L174: I don't quite follow. You say six, but the title says 8. Not directly clear why you say "at the expense of two shrub-PFTs". Probably better to rephrase (this and other parts of the paragraph) by being more exhaustive at first (i.e. there are 6 tree classes and 2 shrubs classes. The reason why is ...)

L185: typo in class 61

L192: Again, the C4/C3 is not clear enough to me, from the text itself. Furthermore, is what Wei et al. did really an accepted way?

L200: you have not clearly stated what the actual "translation" is?

L202: specify that this is only for the European domain considered here

L205: what about C4 crops? According to the authors, the main point of the study is to make a map that reveals traits more than LC, but the C4 vs C3 is perhaps one of the most important ones for vegetation models. So one would expect a proper C4 crop PFT.

L210: more info on what the actual translation means would be needed

L217: Do you mean, with respect to Siebert et al?

L223: I have the impression that the same could have been decided for C4 plants. It seems that the level of scrutiny in this irrigated class is not the same as the one applied for C4/C3. In both cases, an external product could be used, but in C4/C3 it was and not for irrigated.

L229: Are bare classes fully 'non-vegetated', or do they contain residual vegetation? If so, any idea of how much?

L232-236: I do not see what LANDMATE PFT brings specifically here. The distinction between urban and bare is done by the underlying land cover map, so here the credit goes to the LC cover map and not the present study.

L239: what about inland water, i.e. rivers and small lakes? Are these also set as NA? are the RCM land-sea-mask covering such areas?

L282-285: why bring this issue of cost and size here? Here you have a very large amount of records for free... (ie. You do not need to collect them), so these phrases seem irrelevant.

Table 6: I do not understand what these numbers represent exactly. Samples? Correct samples (i.e. where GT-SUR and LANDMATE match)?

Table 7: why not put this as a graph instead of a table? It would be more informative with a single glance

Figure 6 (and others): personally, I find the shadow of the points distracting.

Figure 9: should be "Bucharest"

Format of results is not optimal for understanding. IT should be done in a more condensed way, and not class by class, with a large plot for each. Try to be more synthetic.

L447: Why in LANDMATE PFT only made available for Europe 2015? The ESA CCI LC is available from 1992 to 2020. The transition from LC to PFT seems relatively straightforward. Why would you not produce the time series? The users would surely expect it.

L450: this link does not lead anywhere... at least for me...

L457: "For each ESA-CCI land cover class, an individual CWT is developed...". This does not make sense to me like. The standard way to use the CCI CWT is that it applies weights to each land cover class, and the CWT is the matrix of these weights (rows = LC, cols = PFT). If you apply a CWT to each LC class (to each row), you are just setting the weights for that class. Now I understand that what you are actually doing is to set an extra

dimension, that of the HLZ. So you need to be more specific, and say that the CWT is decomposed by different HLZ for every LC.