Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2020-262-RC2, 2021

Earth System Science Data
Open Access Discussions

Interactive comment on "Development of a global dataset of Wetland Area and Dynamics for Methane Modeling (WAD2M)" *by* Zhen Zhang et al.

Joe Melton (Referee)

joe.melton@canada.ca

Received and published: 12 February 2021

Review of Zhang et al. (ESSD)

Zhang et al. generate a novel dataset of global wetland dynamics using the remotely-sensed dataset, SWAMPS, as a starting point. They then address its shortcomings by bringing in regional datasets for known issues like saturated wetlands in the high latitudes (e.g. using NCSCD) or trouble seeing through canopy in tropics (e.g. using CIFOR). I generally find what they have done to be quite logical. I think the suggestion of an ensemble approach for future versions would be very useful and was glad to see it suggested. At present, the approach here relies quite heavily on the performance of SWAMPS/GLWD/NCSCD/CIFOR being the best products out there. That is unlikely

to be correct for all locations so more regional specific datasets could likely improve future versions of WAD2M. My biggest complaint about the paper was the graphics and specifically the colour schemes. This sounds trivial but it did make interpreting the figures challenging. I hope the authors take my comments there seriously. As I think my comments can be relatively easily dealt with I am suggesting minor revisions.

- line 51: also conversion to rice agriculture?

- l57: does Ramsar include rivers/lakes/ponds as wetlands? That is how I read that but am not sure if that was what was intended.

- l65: Can you provide an example of a dataset that omits them?

- l.91 - detect flooding beneath most vegetation canopies - this begs the question which canopies can it then not detect the flooding underneath?

- l.92 - maybe specify the frequency you are talking about, could be construed to mean satellite return frequency or something rather than frequency in the electomagnetic spectrum, esp as the next sentence talks about temporal coverage.

- l. 104 - Is there any issue with high-latitude regions due to the satellite orbits or the low radiation in the winter that should be mentioned here?

- l. 143 - the problem with this approach is that you then end up potentially increasing the wetland area since you are assuming that SWAMPS is wrong and the wetland product is correct. Given how error prone large-scale wetland mapping can be I think this is likely a dangerous assumption. I don't really disagree with the approach though, as there is not likely a better way to include wetlands that SWAMPS doesn't see. I would suggest that somewhere it is noted that this approach could inflate wetland areas as it would then include wetlands that may be erroneous in the wetland inventory.

- l. 145 - Hugelius has published a new effort for high latitude peatlands, while I understand that this work might have been completed prior to the publication of that work, I wonder if there is much change between the NCSCD than the newer paper (Hugelius

et al. 2020)? I see Hugelius is on the author list.

- l 149 - I wonder if the dataset used for the temperate regions could be improved upon by using updated datasets for more specific regions. E.g. for Canada (Mahdianpari et al. 2020). Since the wetland inventory seems important (my comment for line 143), it seems reasonable to try to get the very best inventories. I am not sure if GLWD is that (although it was a good first attempt).

- l. 181 - what proportion of SWAMPS 3.2. observations were 'valid'? How many valid observations were needed to be able to produce the monthly values?

- l. 183 - What does it mean that they were later 'refiltered'?

- l. 186 - When I tried to use the GSW for a project I found it to have large issues with stripping of the data due to satellite passes and data quality control. Was this problem also encountered and if so how was the influence of the stripping minimized?

- Also with GSW, since it is based on Landsat does this mean it is also just the open waterbodies (i.e. no canopy cover)? I think you say this on line 196 but not earlier in describing what GSW provides so I want to be sure.

- L 211: I think this is getting to my question about how it compares to the newer Hugelius product.

- L. 230 to get an idea of how good the datasets agree for areas of overlap, was any attempt made to see how CIFOR and GLWD compare against each other or NCSCD vs. GLWD? If the agreement is poor then it doesn't allow a lot of confidence in the use of these datasets.

- Table 1 - misplaced comma final column and row.

- It would be useful to explain why SWAMPS was the starting remotely-sensed dataset and not GIEMS. Why one over the other? Assumedly there was some performance improvement, like is seen in Fig 4a where SWAMPS seems more sensitive to high

latitude wetlands?

- In Fig 1 the Inundation box dead-ends. So that information is not integrated into the final output? I think a line is missing. Also isn't the correcting factor applied to dynamic inundation, not becoming the corrected dynamic inundation as is shown? This figure is really hard to follow. I suggest a re-think.

- The maps have a, creative, colour scheme. Have you considered a linear colour map? There are good resource around like, https://sciviscolor.org/. At present there is an arbitrary strong visual cut-off at between 0.1 and 0.2 giving some sort of importance to a threshold that has none. Is the maximum value in all of the dataset 0.6? Also the maps could have Antartica cut off to allow a larger representation of the non-ice covered parts of the world. Also it would be nice if the images could have improved resolution. When I tried to zoom into see the distribution it looks as though the plotting used a setting that interpolated between gridcells, i.e. pixels are washed out. Please check and correct this. BTW, Fig 10 with the number of sources panel is the only time I think the colour scheme makes sense. Here it is categorical so the use of different colours with no continuous colour scheme makes perfect sense (small point on that one - why have a 0 value? Isn't that only for Antarctica which should be trimmed anyway?).

- Fig 5, there is a hashtag you might want to check out #endrainbow. It is based on the fact that rainbow colour scheme tends to distort the perception of the data since values (e.g. yellows) with no greater importance stand out more to readers, but shouldn't as there is nothing important about the 0.5-0.6 fraction. Please consider using colour schemes that follow best practices for data visualization.

- Fig 5 when the grey value is below 0 on the colour bar, how do we interpret that? It is missing/no values or 0? I assume it must be missing/no since Ternetieva isn't really showing wetlands then none for the most S and N part of the domain. But then other plots have the 0/missing/no wetland grey in the middle of their domains, so what is going on?

- L 360 - 'owing to'

- Fig 6 - Sorry to continually complain about colour schemes but why not just use a linear blue colour scale? Why add in the green? It makes me think there is something especially important about middle of the road values, which isn't the case, so please reconsider. E.g. the most visually striking box in the lower plot (WSL) is for WAD2M and GSW - is that really the thing you want to stick in the reader's mind?

- Fig 8 over the same years? What years are they?

- Fig 9 the caption doesn't mention if these are the watershed values and how that might be defined. Might be worth including a small sub-map showing the boundaries.

- l388 - I think the comparison with Parker et al. 2018 is a bit tenuous, wetland extent is only part of what affects CH4 emissions so the consistency could just be fortuitous. Also this confuses me a bit, it says no significant trend for pan-tropical then the next sentence says tropical have net reduction. So how does pan-tropical differ from tropical?

- L430- this sentence seems to be saying two things: 1) further work needed to confirm and 2) improvement over existing. I suggest rewording so it is more one thought rather than a mash of two that conflict.

- L443 - I quite like this point about ensemble maps. I think it would make version 2.0 much better.

- L457 - artificial neural networks are a machine learning approach.

- L461 - what are less informed wetlands?

- Some regions may be 'not wetland' then become wetland for, say a month since that is the shortest time this dataset can do, is the assumption that the system can behave like a wetland in that amount of time? Since the models will run their wetland model over this as a mask I think that is an underlying assumption but I am not sure if it is

true. This is more of a comment than question and I am sure that number of pixels like that are minimal but I think the assumption is made that each time a pixel 'becomes' a wetland it is assumed to immediately adopt a wetland behaviour from a biogeochemical perspective. This is likely out of the scope of this paper but it does play into how this dataset is used.

Refs cited:

Hugelius, G., Loisel, J., Chadburn, S., Jackson, R. B., Jones, M., MacDonald, G., Marushchak, M., Olefeldt, D., Packalen, M., Siewert, M. B., Treat, C., Turetsky, M., Voigt, C. and Yu, Z.: Large stocks of peatland carbon and nitrogen are vulnerable to permafrost thaw, Proc. Natl. Acad. Sci. U. S. A., 117(34), 20438–20446, 2020.

Mahdianpari, M., Salehi, B., Mohammadimanesh, F., Brisco, B., Homayouni, S., Gill, E., DeLancey, E. R. and Bourgeau-Chavez, L.: Big Data for a Big Country: The First Generation of Canadian Wetland Inventory Map at a Spatial Resolution of 10-m Using Sentinel-1 and Sentinel-2 Data on the Google Earth Engine Cloud Computing Platform, Can. J. Remote Sens., 1–19, 2020.

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2020-262, 2020.