

Review of Cui et al. **essd-2021-186**

Anonymous Referee #2

Referee comment on "A 1 km global dataset of historical (1979–2013) and future (2020–2100) Köppen–Geiger climate classification and bioclimatic variables" by Diyang Cui et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-186-RC2>, 2021

General Comments

Cui et al. present a new 1 km Köppen Geiger (KG) historical and future climate datasets as well as associated bioclimatic variables for use in species distribution modeling and other climate and environmental change related applications. They perform a number of sensitivity tests on the classification algorithm design, and demonstrate their new product has a higher accuracy than previous KG products.

The paper is generally well structured and well written, with most of the methods and results clearly described. The authors have included a number of useful analyses that serve as a rigorous assessment of their products. These will be useful new products for the community and definitely merits publication in ESSD. However, I believe the manuscript could benefit from further clarification to the methods, and improvements to some of the results description is needed, as detailed in the specific comments below.

Finally, the authors have clearly made the case for why a 1 km product is needed (as opposed to the much coarser resolutions of most of the existing products) but I think the authors should more clearly state the differences between their product and the other 1 km product that is available (Beck et al., 2018). Please see specific comments.

Specific Comments

Introduction

It would be useful to know how the datasets included and the methods used differ from Beck et al. (2018). I appreciate the authors have already mentioned the potential biases with the Beck et al. paper in the intro, but given it is the only other 1km dataset that they cite I would appreciate knowing more of the differences. I can see some differences in the methods (e.g. line 133) but in lines 185-186 Beck et al. is cited as the method used to select members from the ensemble (therefore a similarity, I presume). A brief few sentences describing the similarities and differences between these two 1km products might be useful in the introduction.

Section 2

Minor point, but it might be useful to know how you selected the datasets to include in your sensitivity analysis. Were all possible datasets at 0.5 degree resolution and finer considered for example?

Were other finer resolution datasets other than CHELSA and WorldClim considered for the downscaling? Not much information is given on how or why WorldClim was used for topographic correction in lines 106-111, or how this differs from line 117-118 where the authors say WorldClim data were used for downscaling? Please could you clarify this in the text?

Table 1 might benefit from an additional column explaining the final use of the dataset (e.g. classification, or "topography correction/downscaling" (in case of WorldClim and CHELSA)).

Lines 81-82: could emphasize again here the use for PFT mapping (as in Poulter et al., 2011; 2015) for ESM simulations.

Section 3

Lines 150-159: do you have a reference for the delta method? I.e. it's not new to this paper?

Lines 184 to 194: So the CHELSA data were not used for downscaling per se, but we used to correct for topography effects by having these data as one ensemble member? Have I understood correctly? Just how the CHELSA data were used could use some clarification.

Section 4

Figures 1 and 3 need to have better resolution and larger text sizes. The text is blurry and/or difficult to read. In figure 3 it is not easy to see the KG classes and even more difficult to see the confidence levels. All figures look like they have been pasted together in another piece of software. I can see lines around the figures in Figure 4.

It would be interesting to discuss why the confidence level changes over time in Fig. 3b. For example, why do you think we suddenly get a pocket of low confidence to the southwest of the great lakes in North America, starting in 1985-2014? I would have thought this would be a high confidence area given the availability of data? Why does the lower confidence region in eastern Russia disappear after 1985?

Line 224: maybe worth calling it an average climate classification map given the climate could have changed over this time period?

Lines 227-228: could the lower confidence levels actually be because the climate has changed over this period, and not just uncertainty in existing climate data?

In Figs. 4 and 5 it might be useful to have an elevation map in a sub figure. Most of the low confidence regions look like they are the high elevation points, with the exception perhaps of a zone over east India and Bangladesh in Fig. 5c? I wonder what is causing that zone of low confidence?

I am a little confused as to why you have overlapping time periods in Figure 6? Eg. 2020-2049 but then 2040 to 2069?

Lines 236-239 and Fig. 6: it would be interesting to have a brief summary of projected changes in KG zones and not just a description of the confidence levels /uncertainties. It is actually hard to see much change in Figure 6a.

Lines 256-257: Might be worth having Figs. 4 and 5 b in Figs. 7 and 8 to facilitate this comparison comparison.

Section 4.3 and Figure 9: What is the overall accuracy and precision of the historical maps? I cannot see this reported anywhere, only the different combinations of datasets in Section 4.4. It might also be worth reporting accuracy and precision ranges of the historical map per region, or another more quantitative description of accuracy and precision results here. The colorbar for the precision in particular is hard to determine what values we are seeing in the maps. It might be good to choose color bars for these figures that have more distinct/different colors.

In Section 4.4 the total accuracy is still using the same validation datasets as in Section 4.3, am I correct? This needs to be a little clearer in the text. , and 4.4 This needs to be made a little clearer in the text.

Table 3: was the Beck et al. (2018) total accuracy calculated in exactly the same way? If not, this is not a good comparison to make. If so (did you calculate this accuracy) then this needs to be described in the methods. Also need a reference to table 3 in lines 283-284. I don't see the Beck et al. accuracy assessment in Figure 10.

The authors speak about validation in describing Figure 10 but this is in Section 4.4 which is entitled "Sensitivity Analysis". Perhaps Sections 4.3 and 4.3 could be combined and the distinction between these results described more clearly?

The authors should consider adding a new sub-section to Section 3 on accuracy assessment, validation, sensitivity analysis etc. For example, why were the GHCN-D and GSOD products chosen as reference datasets? These products, and the reasons for choosing them, nor the methods for validation are not described anywhere in the methods (unless I am mistaken).

Also the methods used to derive Figure 10, including determining the accuracy of the other datasets, is not described in the methods. This is not clear and yet is important for highlighting how this product is better. Descriptions of confidence level, accuracy and precision calculations, sensitivity analysis, validation etc (e.g. lines 265-273, lines 334-336 etc) should go in the methods (Section 3) and it should be clear the different calculations that are done in Section 4.3, 4.4 and 4.6. Accuracy is used in a number of different ways throughout the manuscript. For example, in line 209 the authors mention lower accuracy but I do not think that is the same as the total accuracy presented in

Sections 4.3 and 4.4?

The differences between different KG maps, and how we can use forest cover to assess those maps, is not that clear to me in Figure 11 and the associated text. For example these two statements in line 310-311 and lines 318-319 is not backed up well by figure 11: "Figure 11 illustrate the enhanced regional details of the maps." And "Moreover, the new Köppen-Geiger maps show accurate depiction of important topographic features and correspond closely with tree lines in the forest cover maps over the regions with complex topography (Fig. 11)." A more careful description and/or different presentation of the results in Fig. 11 is needed. Furthermore, in between these two sentences the authors speak about different regions entirely, for which no maps are shown. I think this section needs some improvement.

It might be worth putting the validation of the historical maps before the future maps. I.e. Sections 4.3 – 4.6, which are all based on the historical maps, before 4.2 and 4.7, which are based on the future maps. This would also address my comment about Lines 236-239 and Fig. 6.

Technical Corrections

A number of "the"s missing throughout the text.