

Responses to Referee#1

We thank Referee#1 for the careful review of the manuscript and the constructive comments. We have revised the manuscript based on the comments and provided detailed point-by-point responses to the comments. Our replies are highlighted in blue.

Cui et al. present Köppen-Geiger climate classification maps for ten 30-yr historical periods between 1979-2017 and future periods under different RCP scenarios. The historical maps were derived from combinations of 3 global temperature products and 3 global precipitation products, while the future maps were derived from bias-corrected CMIP5 projections. The authors claimed that their maps can capture recent and future changes in spatial distribution of climate zones. This dataset is useful and relevant for wide audience, but there is an important issue that have to be addressed before publication:

The authors compare two maps for 1980-2009 and 1987-2016 in Fig. 11. However, according to Table 1 and Table 3, the map for 1980-2009 was based on 3 temperature products (CRU, UDEL and CHELSA) and 3 precipitation products (UDEL, CHELSA, GPCC), while the map for 1987-2016 was based on 2 temperature products (CRU and UDEL) and 2 precipitation products (UDEL and GPCC), as CHELSA only covers 1979-2013 (Table 1). Evidenced by Fig. 5, the CHELSA precipitation have a large impact on the KGC map. It is also be seen in Fig. 14, where the abrupt changes are all happening between 1983-2012 and 1984-2013. So I suspect the difference for the two periods largely come from inconsistent inputs for the two periods, and is not a reflection of the true shift in climate zones. I strongly recommend the authors to discuss the impact of having CHELSA before 2013 and not having CHELSA afterwards on their time series of KGC map.

Thank you for pointing out the issue caused by the inconsistent data inputs for the historical periods. To address it, we excluded the KGC maps based on the 2 temperature products (CRU and UDEL) and 2 precipitation products (UDEL and GPCC) and kept the KGC maps based on 3 temperature products (CRU, UDEL and CHELSA) and 3 precipitation products (UDEL, CHELSA, GPCC) as our final climate map product. According to the evaluation results, the final map product based on CHELSA, CRU, GPCC and UDEL data exhibits higher overall classification accuracy than the maps based on CRU, GPCC and UDEL data. Additionally, the CHELSEA precipitation data considered topo-climatic drivers and predicted better precipitation patterns (Karger et al., 2017), which may not be well captured in statistically downscaled 0.5° climatology data. Having CHELSA as data as input of climate maps can slightly improve the overall classification accuracy and present more accurate depiction of spatial patterns. Therefore, we decided to use the KGC maps based on CHELSA, CRU, UDEL and GPCC as our final climate map product. But since the CHELSA climatology data has a temporal span from 1979-2013, the final KGC map product will cover six historical periods, including 1979-2008, 1980-2009, 1981-2010, 1982-2011, 1983-2013, and 1984-2013. We have updated it in the revised manuscript and created a new version of the historical KGClim dataset.

There are also some minor issues:

1. A grammar mistake in L17: “The new maps offer higher classification accuracy”, higher than which products?

Thank you for this comment. Based on our validation results, our new historical climate maps demonstrate higher overall classification accuracy than all the other existing climate map

products. We have added the detail and revised the sentence to “The new maps offer higher classification accuracy than existing climate map products”.

2. L30 and L36: the authors repeated the definition of the Köppen classification in these two lines, but used “annual cycles” in L30 and “seasonal cycles” in L36. What are the difference?

Thank you for pointing out the issue of the use of “annual cycles” and “seasonal cycles” regarding the definition of the Köppen classification. The Köppen classification is defined based on the seasonal phase of temperature and precipitation annual cycles. Using “annual cycles” and “seasonal cycles”, we refer to the seasonality of temperature and precipitation. To clarify the difference, we have changed “annual cycles” to “seasonal phase of annual cycles”.

3. L103-L104: “Evaluation results indicated that incorporating only CRU, UDEL temperature datasets and UDEL, GPCC precipitation datasets led to higher accuracy in the classification results.” Table 3 tells me the combinations of CRU, UDEL and CHELSA temperature and UDEL, GPCC and CHELSA precipitation lead to the highest accuracy. So I don’t understand why this sentence here did not mention CHELSA.

Thanks for your comment. To decide which datasets to use to generate the historical KGC maps, we first tested several observational climatology datasets with coarse resolution of 0.5°, including CRU, UDEL, GHCN_CAMS, GPCC, UDEL, and PREC/L. We found out that using only CRU, UDEL temperature datasets, and the UDEL, GPCC precipitation datasets have better classification results than having all the CRU, UDEL, GHCN_CAMS, GPCC, UDEL, GHCN_CAMS and PREC/L datasets as input. To clarify this, we have revised the sentence, “Evaluation results indicated that incorporating only CRU, UDEL temperature datasets and UDEL, GPCC precipitation datasets and excluding GHCN_CAMS and PREC/L datasets led to higher accuracy in the classification results”. The paragraph following the sentence introduced the CHELSA dataset and explained the reasons for using the 1-km CHELSA and WordClim datasets in addition to the 0.5° datasets to correct topographic effects and provide better description of precipitation patterns. Table 3 shows the accuracy results with or without CHELSA dataset as input or using CHELSA data alone.

4. L120: Please give the ref.

Thanks for this comment. We have provided the reference, “Technical evaluation showed that the bias-correction method that CCAFS data applied reduced climate model bias by 50–70%, which could potentially address the bias issue in model simulations for the threshold-based Köppen classification scheme (Navarro-Racines, Tarapues, Thornton, Jarvis, & Ramirez-Villegas, 2020).”

5. L123: The Köppen climate classification scheme was first introduced in 1884 (Rubel, F. & Kottek, M. Comments on: The thermal zones of the Earth by Wladimir Köppen (1884). (2011) doi:10.1127/0941-2948/2011/0258.)

We have corrected the year to 1884 and added the reference in the revised manuscript.

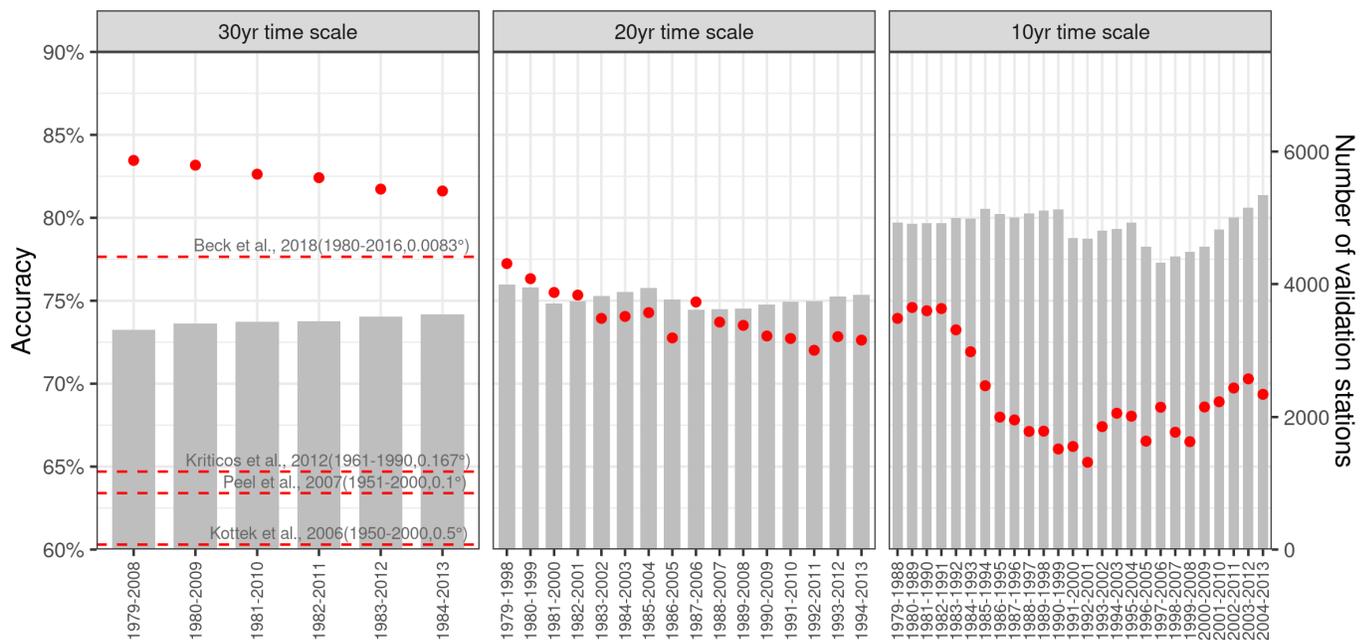
Köppen, W. (1884). Die Wärmezonen der Erde, nach der Dauer der heissen, gemässigten und kalten Zeit und nach der Wirkung der Wärme auf die organische Welt betrachtet. *Meteorologische Zeitschrift*, 1(21), 5–226.

6. L125: KGC is not explained for the first time being used.

Thanks for pointing it out. We have added the explanation of KGC for its first time mentioned in the manuscript.

7. 10. The accuracy of Beck et al. (2018) not plotted.

We have added the accuracy of Beck et al. (2018) in Figure 10.



8. L288: “Duplicate stations in the two datasets were further removed.” This sentence should be moved to section 4.3?

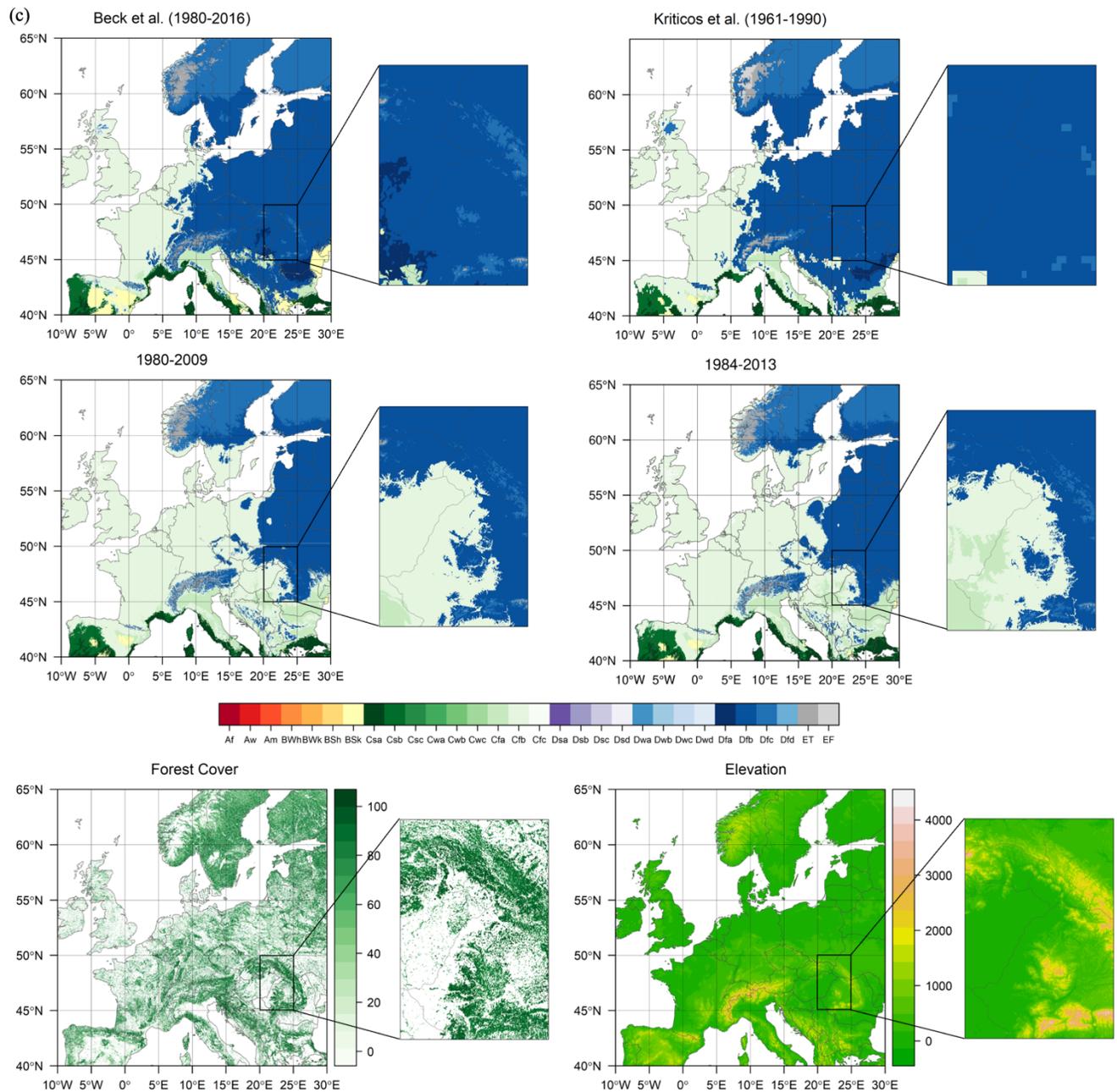
We accepted the suggestion and moved the sentence to section 4.3.

9. L306-307: If the products here are better than previous one, how the previous “worse” maps can be used for “evaluation”?

Thanks for raising the concern regarding the evaluation of the map product. We used the previous climate maps in addition to forest cover and elevation maps for regional and continental scale comparison to identify the potential improvements of the map product. The expression of “evaluation” is not appropriate and we have changed the sentence to “We compared the new Köppen-Geiger climate classification maps with the high-resolution Köppen-Geiger maps from two previous studies, Beck et al., (2018), and Kriticos et al., (2012).”

10. L315-317: “Another improvement ..., which show better agreement with global boreal forest distributions”. I do not find the evidence from the figure.

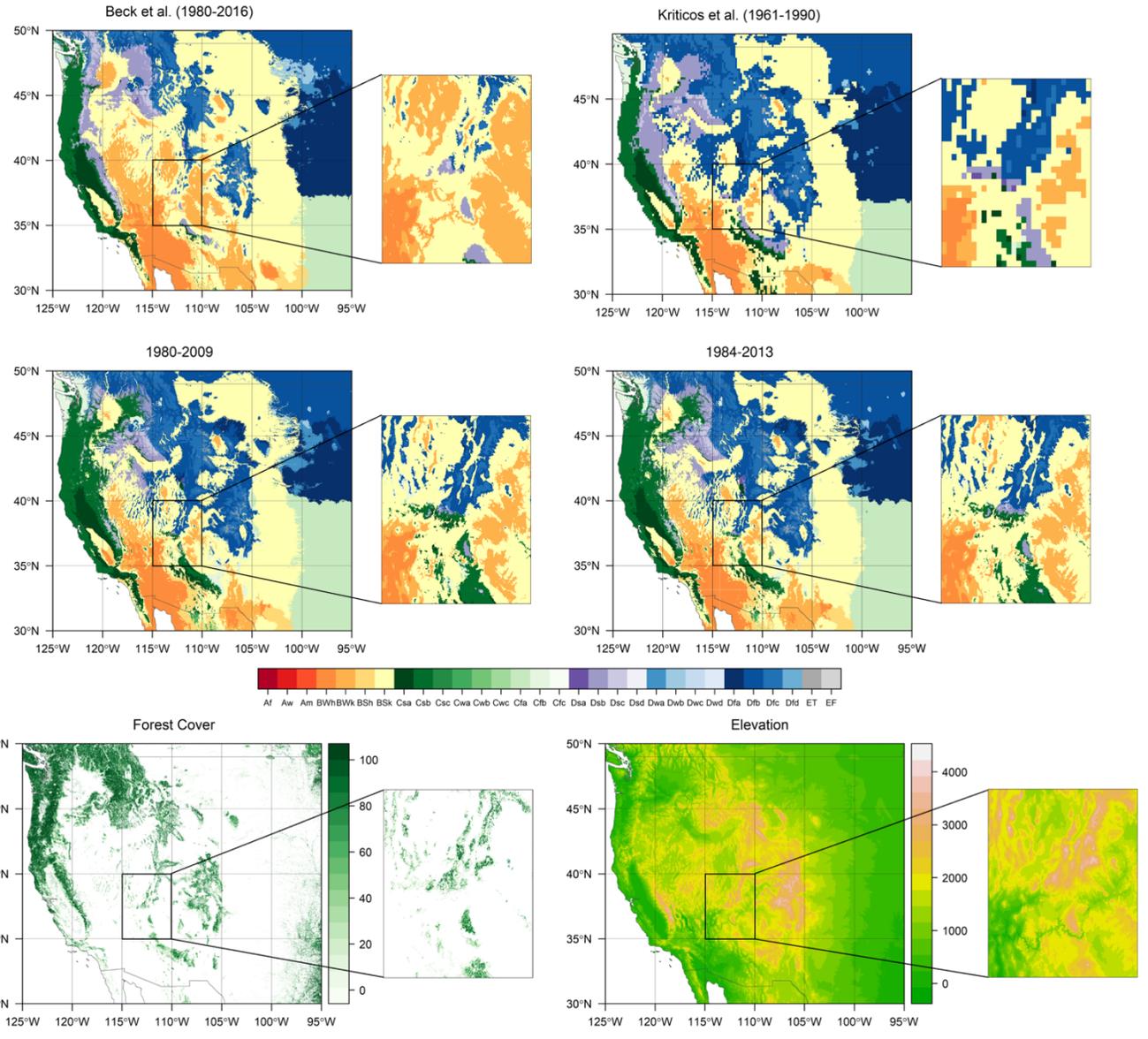
Thank you for your comment. We updated Figure 11 to show the continental and regional scale comparison with more details of the Köppen-Geiger climate classification maps from previous studies, Beck et al., 2018 (1-km, 1980-2016), and Kriticos et al., 2012 (0.167°, 1961-1990), our study (1-km, 1979-2009 to 1984-2013), forest cover and elevation maps. Shown in the updated figure below for Europe, we can see that the new Köppen-Geiger climate classification maps from our study show better agreement with the boreal forest in Carpathian Mountains across Central and Eastern Europe at small scale. The reason causing the different boreal (D) climate zone distribution compared with Beck et al., (2018), and Kriticos et al., (2012), is that we followed the Köppen-Geiger climate classification as described in Kottek, Grieser, Beck, Rudolf, and Rubel (2006), and Rubel and Kottek (2010), and used the threshold of -3 °C as the boundary of temperate (C) and boreal (D) climate zones while Beck et al., (2018), and Kriticos et al., (2012) used the Russell’s modification (1931), which is based the distributions of topographical features and vegetation in western United States. We have added the explanation and description of the figure in the revised manuscript.



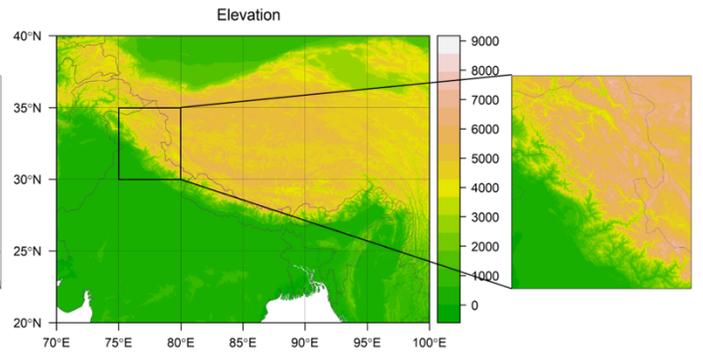
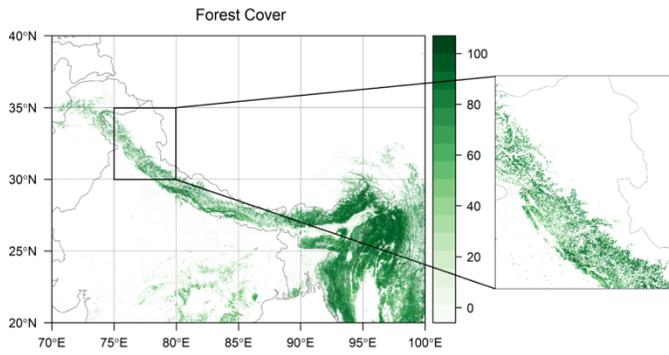
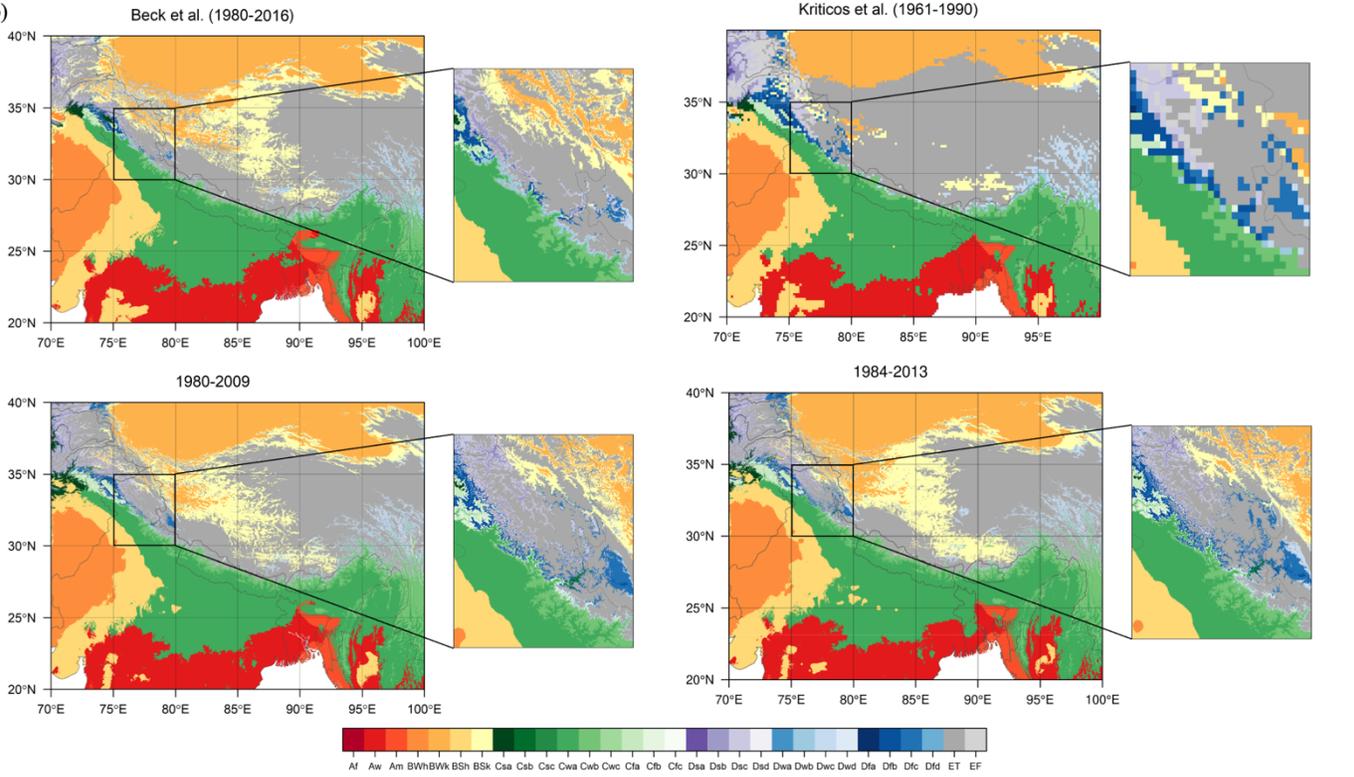
11. L318-L319: “Moreover, the new Köppen-Geiger maps show accurate depiction of important topographic features and correspond closely with tree lines in the forest cover maps over the regions with complex topography”. I cannot see from Fig 11 why the new maps are better than Beck et al., (2018) and Kriticos et al., (2012).

Thanks for raising the concern about Figure 11. We updated Figure 11 to show map comparisons at both continental and regional scale with more enhanced details. We also included more descriptions of the map comparison in the Figure. Here are the updated figure and the revised version of the section:

(a)

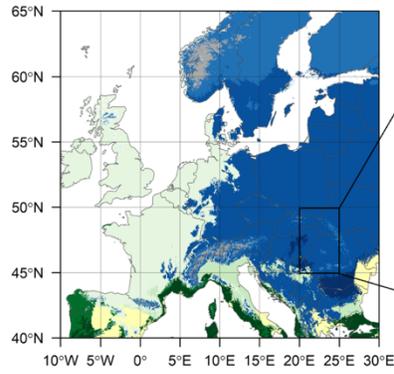


(b)

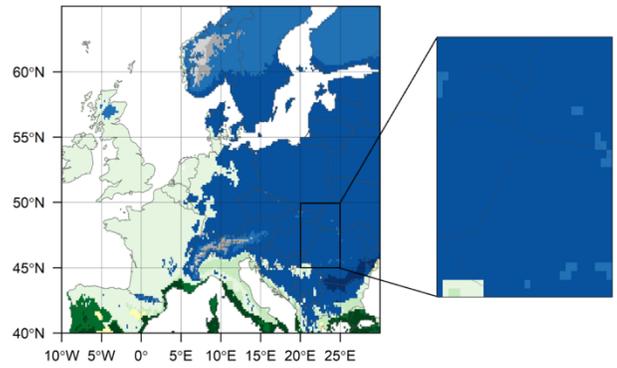


(c)

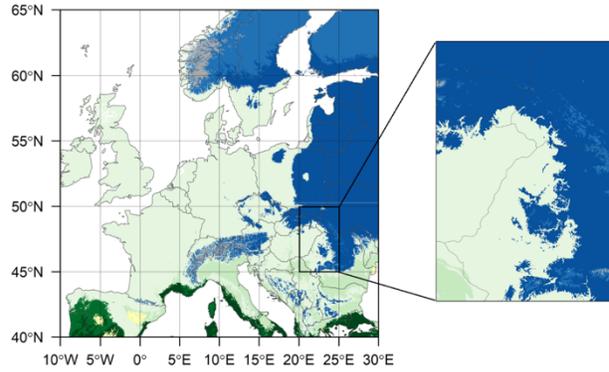
Beck et al. (1980-2016)



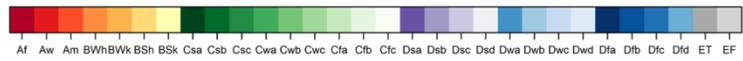
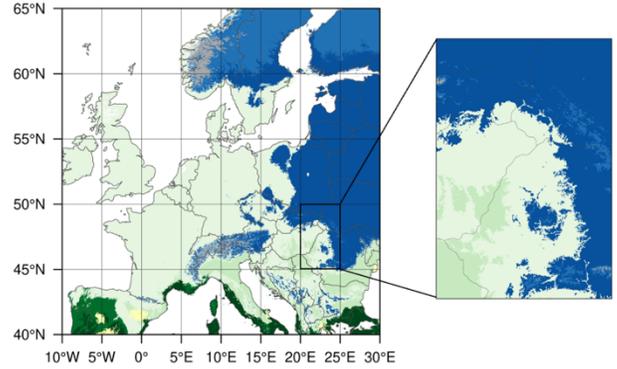
Kriticos et al. (1961-1990)



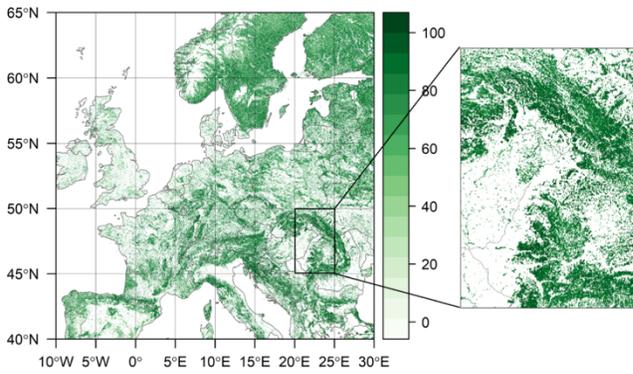
1980-2009



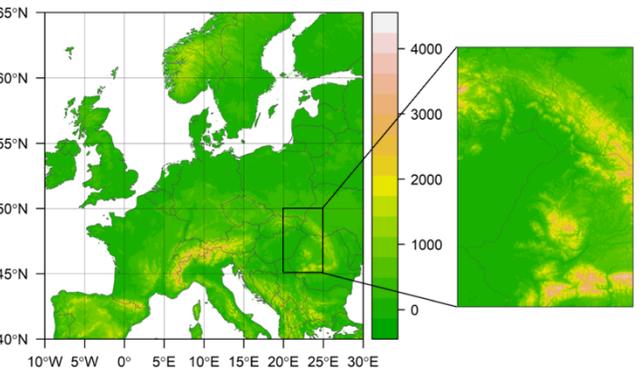
1984-2013

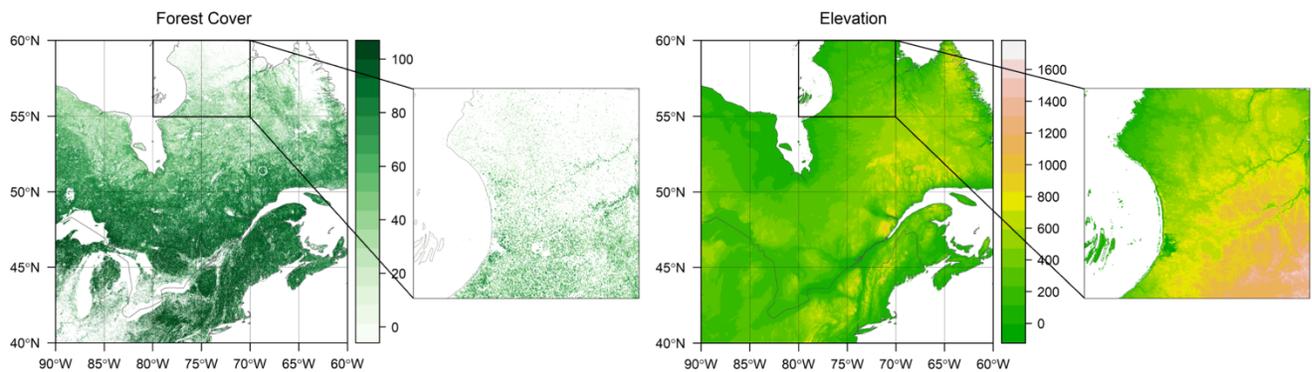
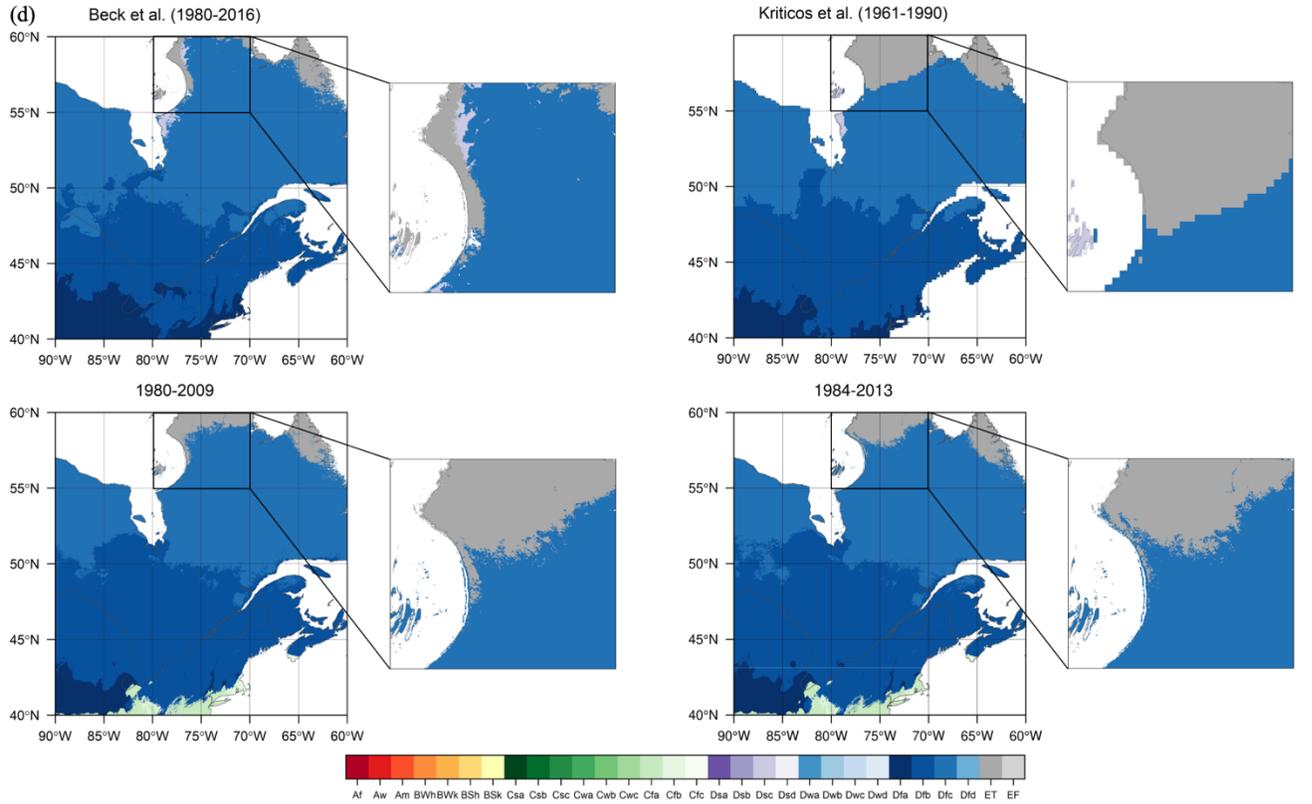


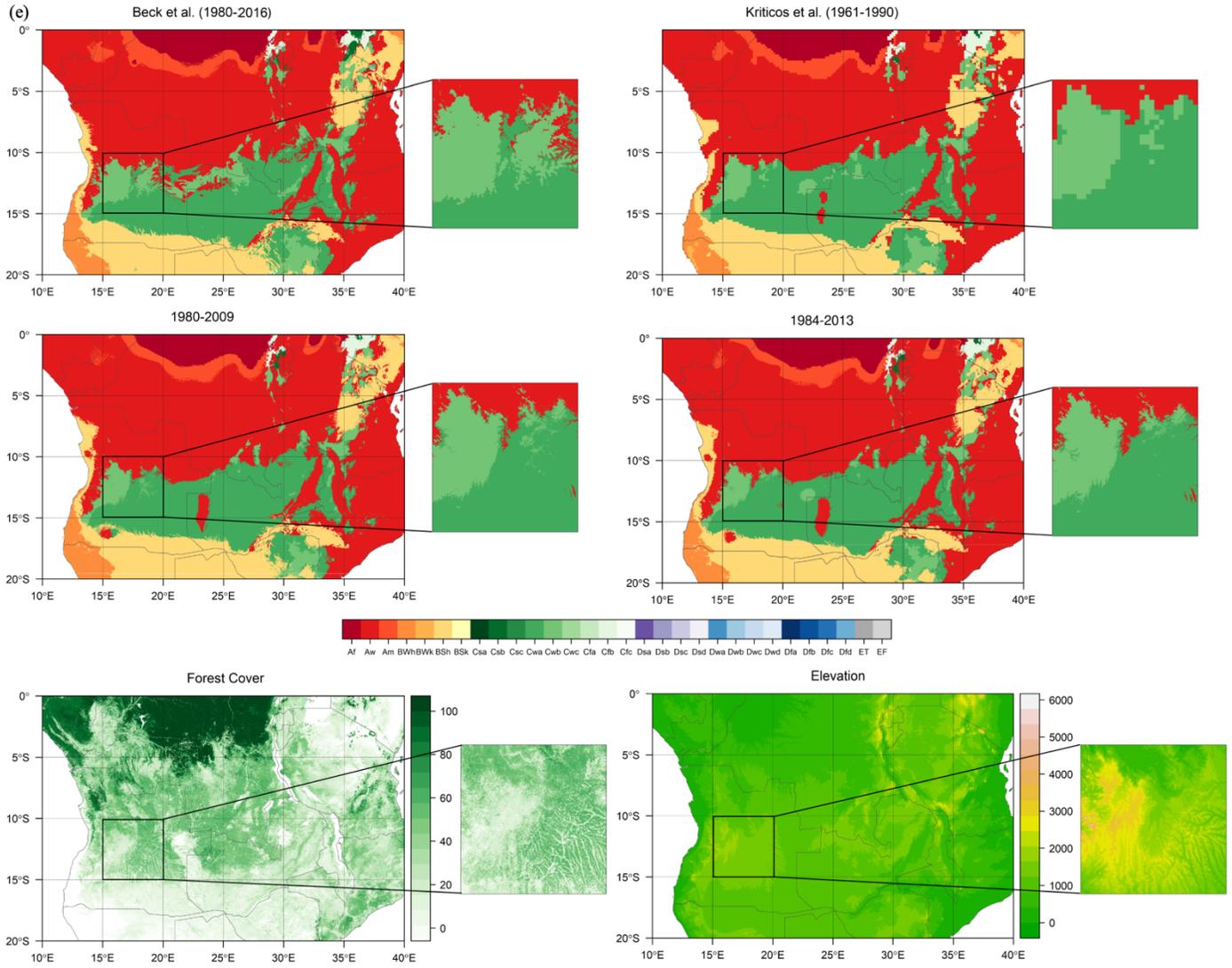
Forest Cover



Elevation







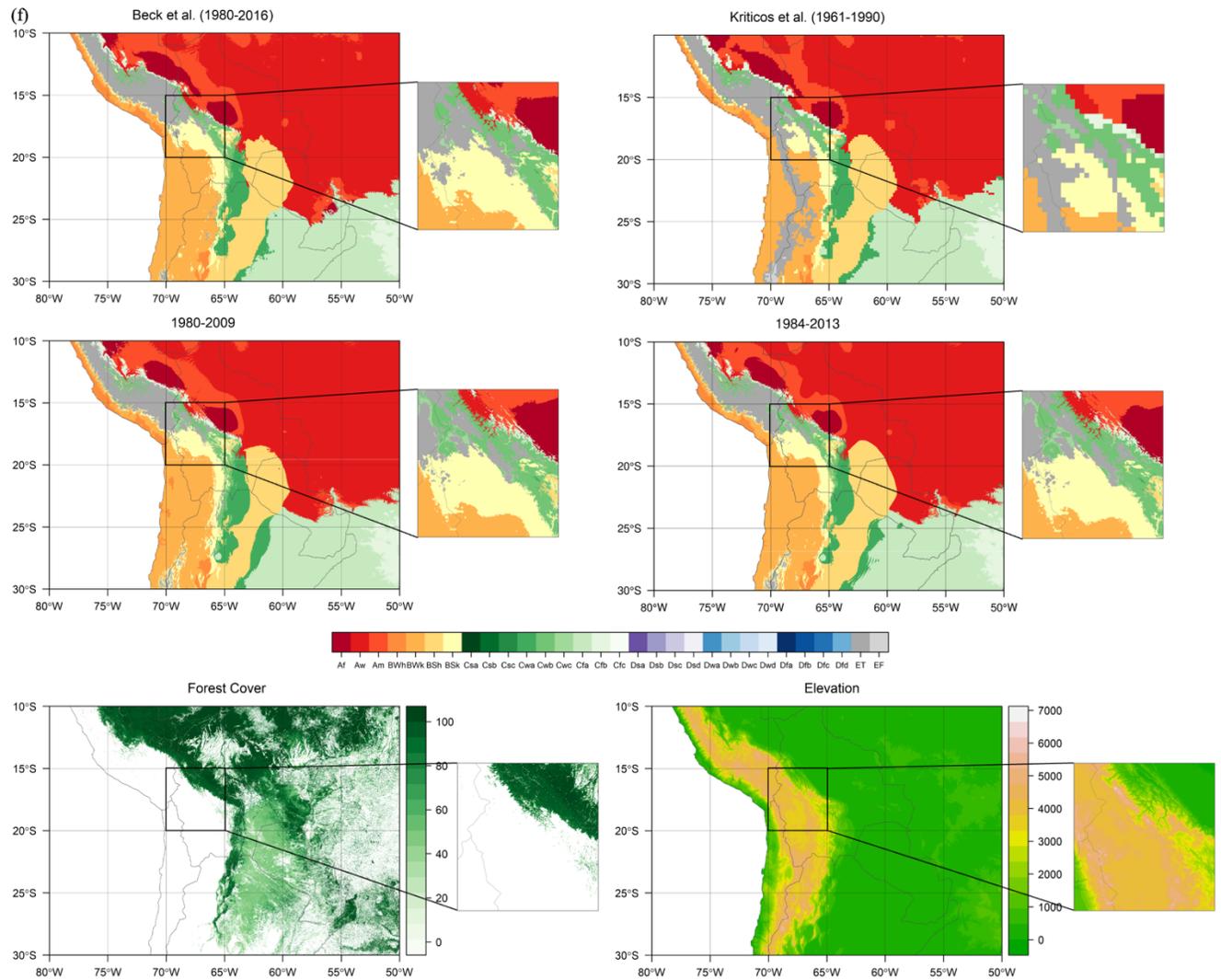


Figure 8. Köppen-Geiger climate classification maps from previous studies, Beck et al., 2018 (1-km, 1980-2016), and Kriticos et al., 2012 (0.167°, 1961-1990), our study (1-km, 1979-2009 to 1984-2013), associated forest cover and elevation maps, for regions with large spatial gradients in climates or sharp elevation gradients. (a) central Rocky Mountains, (b) Tibetan Plateau, (c) Europe, (d) high latitudes in North America, (e) Central and eastern Africa, and (f) central Andes. The forest cover map is the 30m Landsat-based forest cover map for year 2000 (Hansen et al., 2013). The elevation data is the NASA SRTM Digital Elevation 30m data (Farr et al., 2007). The representative period of each map is listed in parentheses.

“At the regional and continental scale, we compared our Köppen-Geiger climate classification maps with previous map products for regions with large spatial gradients in climates, including central and eastern Africa, Europe, North America, and regions with sharp elevation gradients, including Tibetan Plateau, central Rocky Mountains, central Andes (Fig. 8). We compared the new 1-km Köppen-Geiger climate classification maps from our study for time periods of 1980-2009, and 1984-2013 with the high-resolution Köppen-Geiger maps from two previous studies, Beck et al., (2018), which has a resolution of 1-km and temporal coverage of 1980-2016, and Kriticos et al., (2012), which has a resolution of 0.0167° and covers 1961-1990. The Köppen classifications demonstrate good correlation with natural landscape distributions (Belda, Holtanová, Halenka, & Kalvová, 2014; Köppen, 1936; Trewartha, 1954). To show the agreement between the improved Köppen-Geiger

climate classification maps and regional landscape distributions, we also showed maps of forest cover, and elevation distribution for these regions. Figure 8 illustrate the enhanced regional details of the maps.

Compared with the Köppen-Geiger climate maps from previous studies with only one time period, the series of the Köppen-Geiger climate maps from our study demonstrate the ability to capture recent changes in spatial distributions of climate zones. For example, our maps can detect the significant changes in the climate zones specifically driven by the accelerated global warming since the 1980s, for example, the poleward movements of boreal (D) and polar (E) climates in high latitudes in North America shown in the comparison between the 1980-2009 and 1984-2013 Köppen-Geiger climate maps (Fig. 8d). Another example is the expansion of savanna (Aw) climate into temperature (Cw) climate zone, witnessed in Central Africa (Fig. 8e).

Another improvement of the new series of the Köppen-Geiger climate maps is the application of threshold of -3 oC as the boundary of temperate (C) and boreal (D) climate zones, which show better agreement with global boreal forest distributions at regional scale compared with Russell's modification of 0 oC (1931), which Beck et al., (2018), and Kriticos et al., (2012) utilized. Based on the comparison results of the Köppen climate zones and the biome classifications from the World Wildlife Federation (Rohli, Joyner, Reynolds, & Ballinger, 2015), the boreal (D) climate zone largely corresponds to the distribution of boreal forest (Cui, Liang, & Wang, 2021). For example, evidenced in Figure 8c, the new Köppen-Geiger climate classification maps from our study show better agreement with the boreal forest in Carpathian Mountains across Central and Eastern Europe than Beck et al., (2018), and Kriticos et al., (2012). Figure 8d also shows good agreement of the northern boundary of boreal (D) climate zone in northern part of Quebec in Canada with the boundary of Canada's boreal forest.

Moreover, the new Köppen-Geiger maps can show accurate depiction of important topographic features over the regions with complex topography. For example, the topoclimate of the Himalays southern front determined by the mountain ranges are represented with more details in the new Köppen-Geiger maps compared with Beck et al., (2018), and Kriticos et al., (2012) (Fig. 8b). The abrupt changes in climate along the edges of the Andes mountains are also well described in the new maps (Fig. 8f).

In addition, the distribution of tropical (A), temperate (C) and boreal(D) climate zones in the new Köppen-Geiger maps correspond closely with tree lines in the forest cover maps. The temperate (C) and boreal(D) climate distributions based on the Köppen-Geiger maps show a better agreement with the forest distributions of the Middle and Southern Rocky Mountains than Beck et al., (2018), and Kriticos et al., (2012) (Fig. 8a). For another example, the boundaries of the tropical rainforest in Central Africa and South America are clearly delineated in the in the new Köppen-Geiger maps (Fig. 8e and 8f)."

Responses to Referee#2

We thank Referee#2 for the careful review of the manuscript and the constructive comments. We have revised the manuscript based on the comments and provided detailed point-by-point responses to the comments. Our replies are highlighted in blue.

General Comments

Cui et al. present a new 1km Koppen 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 1km 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 1km 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.

Thanks for your suggestion. We have added more detailed description of the 1-km climate map product from Beck et al. (2018) in Introduction:

“As the only 1-km global climate classification map product, Beck et al. (2018) provided global climate classification maps for two periods 1980-2016 and 2071-2100 under RCP8.5. The maps were derived using climate data from WorldClim V1 and V2 (Fick & Hijmans, 2017), CHELSA V1.2 (Karger et al., 2017), and CHPclim V1 (Funk et al., 2015). To represent historical climates, they adjusted the inconsistent temporal spans of climatology

datasets to the period 1980-2016, by adding interpolated temperature change offsets or multiplying precipitation factors, which may lead to biased coverage of the historical period.”

The major differences between our product and Beck et al. are the climatology datasets used, downscaling approach, temporal coverage, and the definition threshold for temperate (C) and boreal (D) climate zones. For similarities, we used the same highest agreement method to determine the final climate class from an ensemble of climate maps. We have explained the details about our product in comparison with Beck et al. in the revised manuscript:

“In this study, we presented an improved long-term Köppen-Geiger climate classification map series for 1) six historical 30-yr periods of the observational record (1979-2008, 1980-2009, 1981-2010, 1982-2011, 1983-2012, 1984-2013) and four future 30-yr periods (2020-2049, 2040-2069, 2060-2089, 2070-2099) under four RCPs (RCP2.6, 4.5, 6.0 and 8.5). To improve the classification accuracy and achieve a resolution as fine as 1-km (30 arc-second), we combined multiple datasets, including WorldClim V2 (Fick & Hijmans, 2017), CHELSA V1.2 (Karger et al., 2017), CRU TS v4.03 (New, Hulme, & Jones, 2000), UDEL (Willmott & Matsuura, 2001), GPCP datasets (Beck, Grieser, & Rudolf, 2005) and bias-corrected downscaled Coupled Model Intercomparison Project Phase 5 (CMIP5) model simulations (Navarro-Racines et al., 2020) (Table 1). We used the WorldClim Historical Climate Data V2 (Fick & Hijmans, 2017) to downscale the 0.5° climatology datasets including CRU, UDEL and GPCP, and derive high resolution climate data for the historical periods. To determine the final climate class, we used the climate class with the highest agreement level from an ensemble of climate maps derived from different combinations of surface air temperature and precipitation products, as implemented in Beck et al. (2018).”

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?

Thank you for raising the concern about the dataset selection. To decide which datasets to use to generate the historical climate maps, we first reviewed the climate datasets used in previous climate map products (Cui, Liang, & Wang, 2021). We considered all the datasets at 0.5° and finer resolution and selected the datasets with a global spatial coverage, temporal span longer than 30 years, and spatial resolution of 0.5°. These datasets include CHELSA, CRU, UDEL, GHCN_CAMS, GPCP, UDEL, and PREC/L. We focused on the gauge-based climatology datasets because generally they provide long-term records and better agreement with ground observations compared with other satellite-based, and re-analysis datasets (Sun et al., 2018). Then we tested several combinations of the observational climatology datasets and found out that using the CHELSA, CRU, UDEL temperature datasets, and the CHELSA, UDEL, GPCP precipitation datasets and excluding GHCN_CAMS and PREC/L datasets can lead to better classification results.

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?

Thanks for the comment. We selected CHELSA and WorldClim datasets as they are the 1km climate datasets with global spatial coverage, long-term temporal coverage, and have been widely applied in climate and ecological studies. We used the WorldClim data to downscale the 0.5° climatology data to derive high resolution climate data with longer and more recent temporal coverage. It is because the WorldClim dataset which has a temporal span of 1970-2000, cannot cover the time periods of the present KGC maps from 1979 to 2013. In addition, the WorldClim dataset used topo-climate factors, including elevation, distance to coast and satellite-derived covariates to interpolate weather station data to 1 km resolution. The delta approach using WorldClim data as reference data to downscale the at 0.5° climatology data provides a simple and easy way to correct topographic effects and reduce potential bias in the downscaled climate data.

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).

Thanks for your suggestion. We have added a column to explain the use of the climatology datasets in Table 1.

Table 1 Climatology datasets to generate present global maps of Köppen climate classification with varied spatial resolutions

Dataset	Usage	Spatial Res.	Temporal Span	Variable	Source and Description
Present Köppen classification map series with resolution of 30 arc-second (1km)					
CRU	Map Input	0.5°	1979-2017	T	Climatic Research Unit (CRU) TS v4.03
UDEL	Map Input	0.5°	1979-2017	T, P	U. of Delaware Precipitation and Air Temperature
WorldClim	Downscaling	0.0083°	1970-2000	T, P	WorldClim Historical Climate Data V2
CHELSA	Map Input	0.0083°	1979-2013	T, P	Climatologies at high resolution for the earth's land surface areas (CHELSA)
GPCC	Map Input	0.5°	1979-2016	P	Global Precipitation Climatology Centre (GPCC)
PREC/L	Data Selection	0.5°	1979-2012	P	NOAA's PRECipitation REConstruction over Land (PREC/L)
GHCN_CAMS	Data Selection	0.5°	1979-2017	T	GHCN_CAMS Gridded 2m Temperature (Land)
Future Köppen classification map series with resolution of 30 arc-second (1km)					
CMIP5	Map Input	0.0083°	2020-2100	T, P	CCAFS-Climate Statistically Downscaled Delta Method CMIP5 data
WorldClim	Downscaling	0.0083°	1970-2000	T, P	WorldClim Historical Climate Data V2

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

We accepted the suggestion. We added the emphasis of the application of KGC maps for plant functional type mapping 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?

Thanks for your comment. We cited the reference which introduced the delta method and used delta change to downscale model simulations in Section 3.2:

“Due to limited number of available observational datasets with high resolution and long-term continuous temporal coverage, the research implemented the delta method by applying a delta change or change factor (Hay, Wilby, & Leavesley, 2000; Wilby & Wigley, 1997) onto the WorldClim historical observations (Fick & Hijmans, 2017) to achieve 30-yr average climatology data with a 1-km resolution based on the CRU, UDEL and GPCC datasets. The delta method is a statistical downscaling method that assumes that the relationship between climatic variables remain relatively constant at local scale (Wilby & Wigley, 1997).”

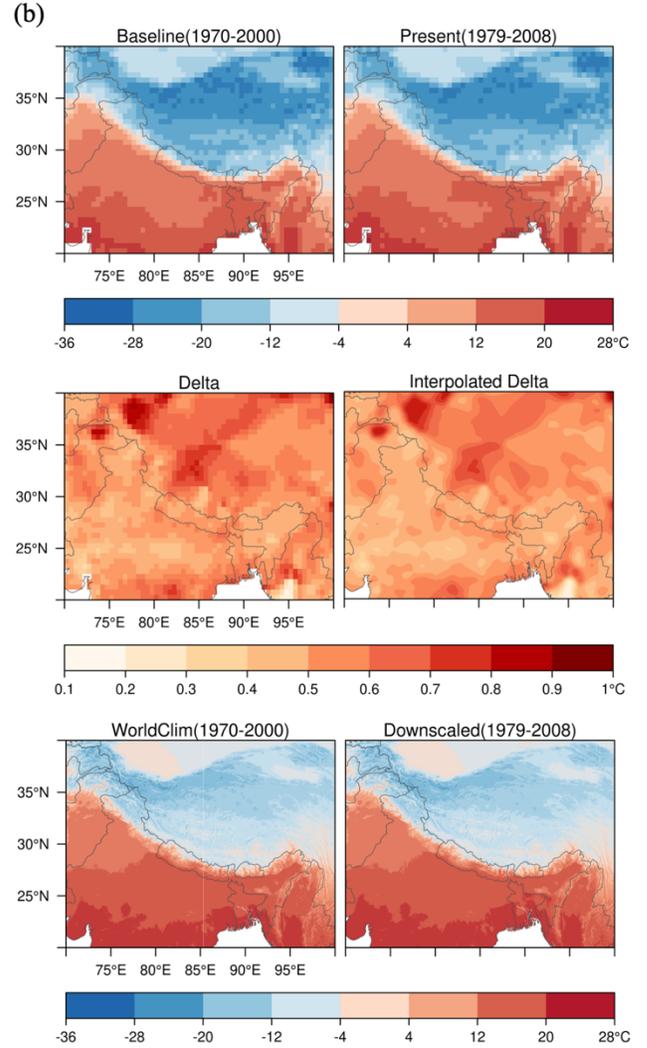
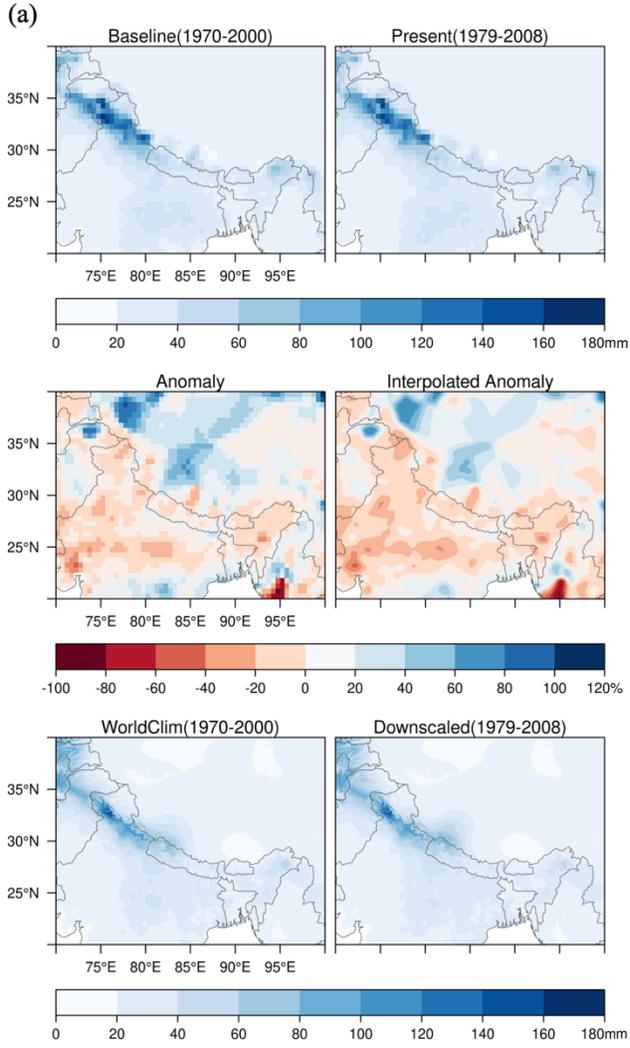
Lines 184 to 194: So the CHELSA data were not used for downscaling per se, but we used to correct for topography effects by having these data as one ensemble member? Have I understood correctly? Just how the CHELSA data were used could use some clarification.

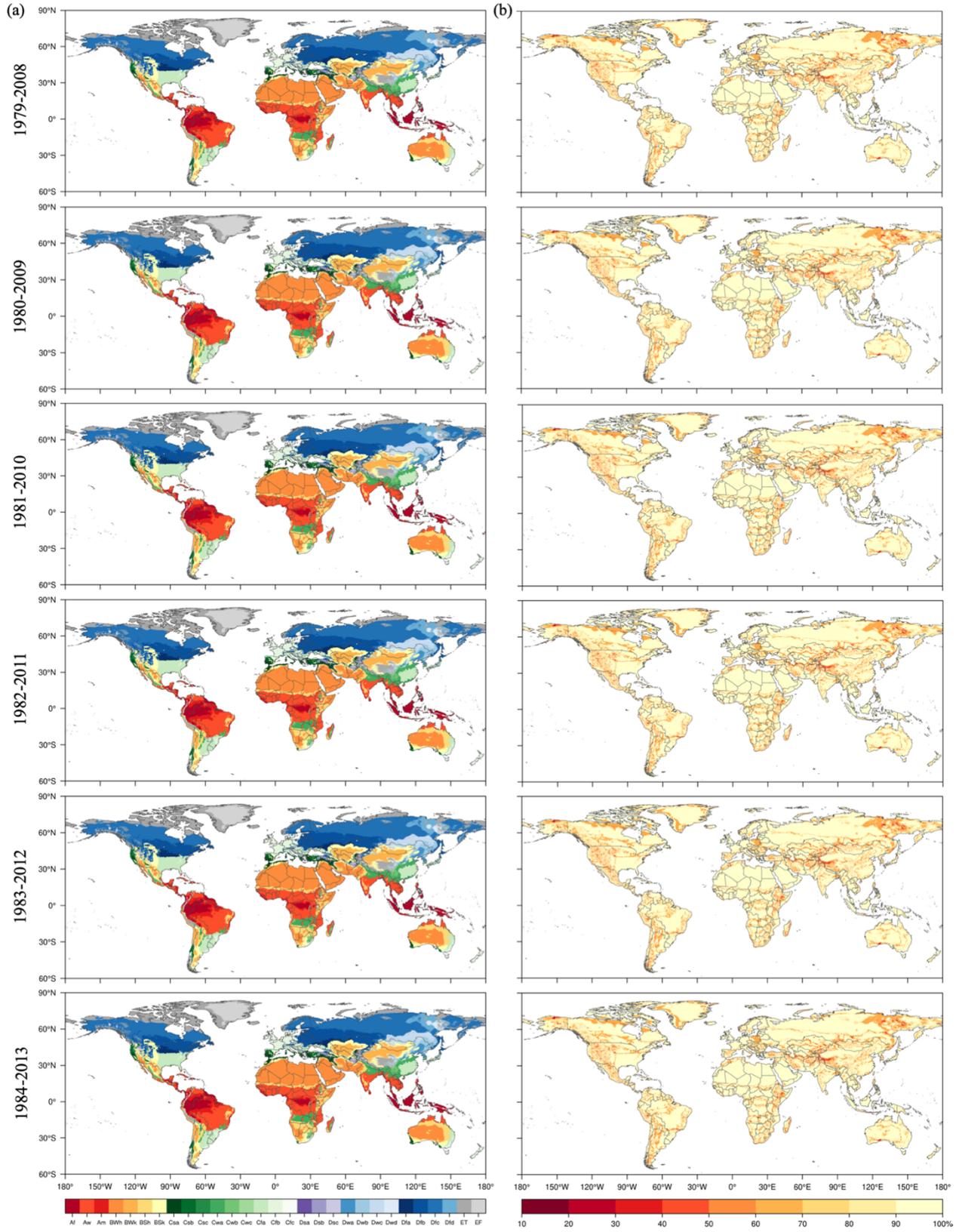
Correct. The reason is that different from the WorldClim data, the 1-km CHELSA dataset covers long-term period from 1979 to 2013, which can be directly used as one ensemble member to generate historical KGC map series. The CHELSA precipitation data also considered topo-climatic drivers and can predict better precipitation patterns (Karger et al., 2017), which may not be well captured in statistically downscaled 0.5° climatology data.

Section 4

Figures 1 and 3 need to have better resolution and larger text sites. 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.

Thank you for your feedback. We have updated figure 1 and 3 with better resolution and larger texts and fixed the line issue in figure 4. Please see the attached updated figures for details.





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?

Thank you for raising the concern about the confidence level changes. It is because that the KGC maps starting from 1985-2014 are based on the 2 temperature products (CRU and UDEL) and 2 precipitation products (UDEL and GPCC) and the maps before 1985-2014 are based on 3 temperature products (CRU, UDEL and CHELSA) and 3 precipitation products (UDEL, CHELSA, GPCC). The CHELSA dataset has a temporal span from 1979-2013, which is not used to generate the KGC maps after 1985-2014.

To address the issue of inconsistent data input, we decided to keep the KGC maps based on CHELSA, CRU, GPCC and UDEL data as our final climate map product. According to the evaluation results, the final map product based on CHELSA, CRU, GPCC and UDEL data exhibits higher overall classification accuracy than the maps based on CRU, GPCC and UDEL data. Additionally, the CHELSA precipitation data considered topo-climatic drivers and predicted better precipitation patterns (Karger et al., 2017), which may not be well captured in statistically downscaled 0.5° climatology data. Having CHELSA as data as input of climate maps can slightly improve the overall classification accuracy and present more accurate depiction of spatial patterns. Therefore, we decided to use the KGC maps based on CHELSA, CRU, UDEL and GPCC as our final climate map product. But since the CHELSA climatology data has a temporal span from 1979-2013, the final KGC map product will cover six historical periods, including 1979-2008, 1980-2009, 1981-2010, 1982-2011, 1983-2013, and 1984-2013. We have updated Figure 3 in the revised manuscript and the patterns of KG climate and corresponding confidence levels show good consistency in the updated Figure 3.

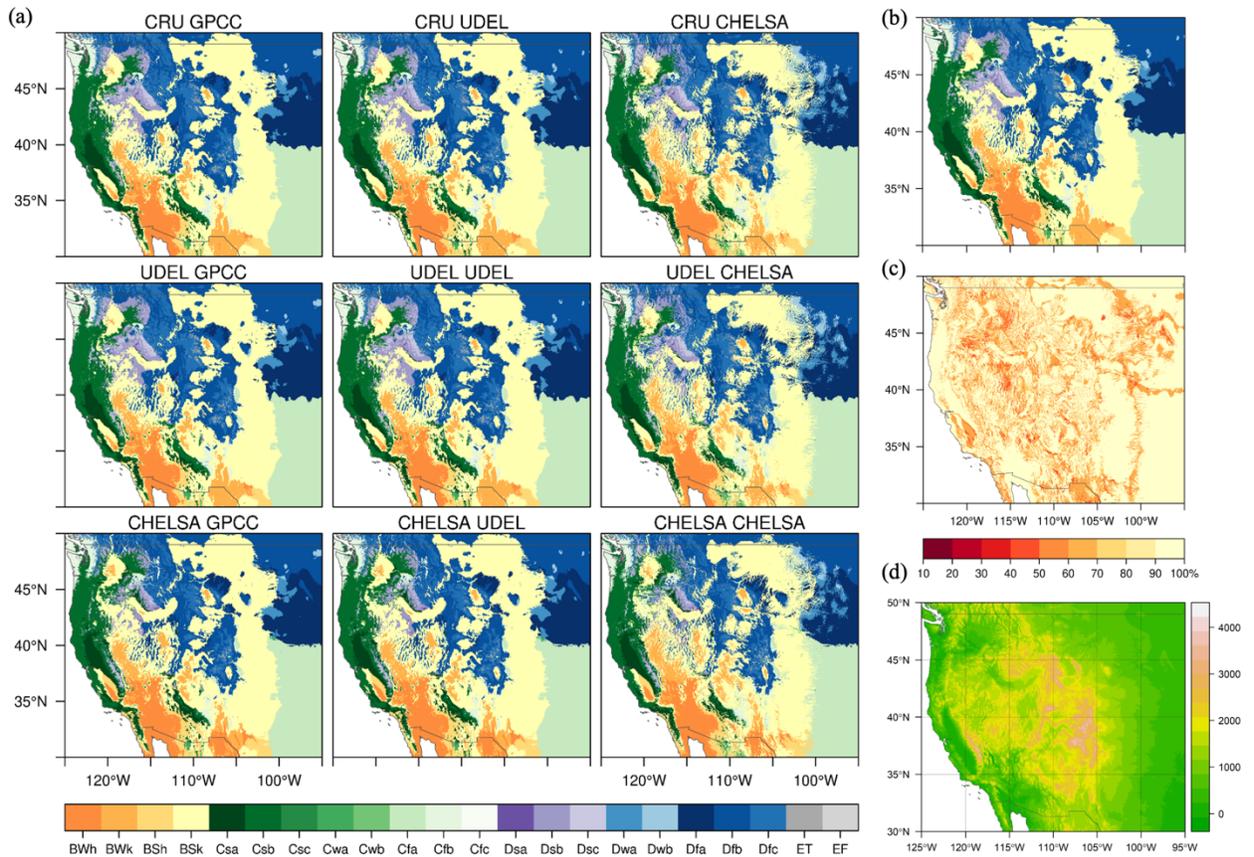
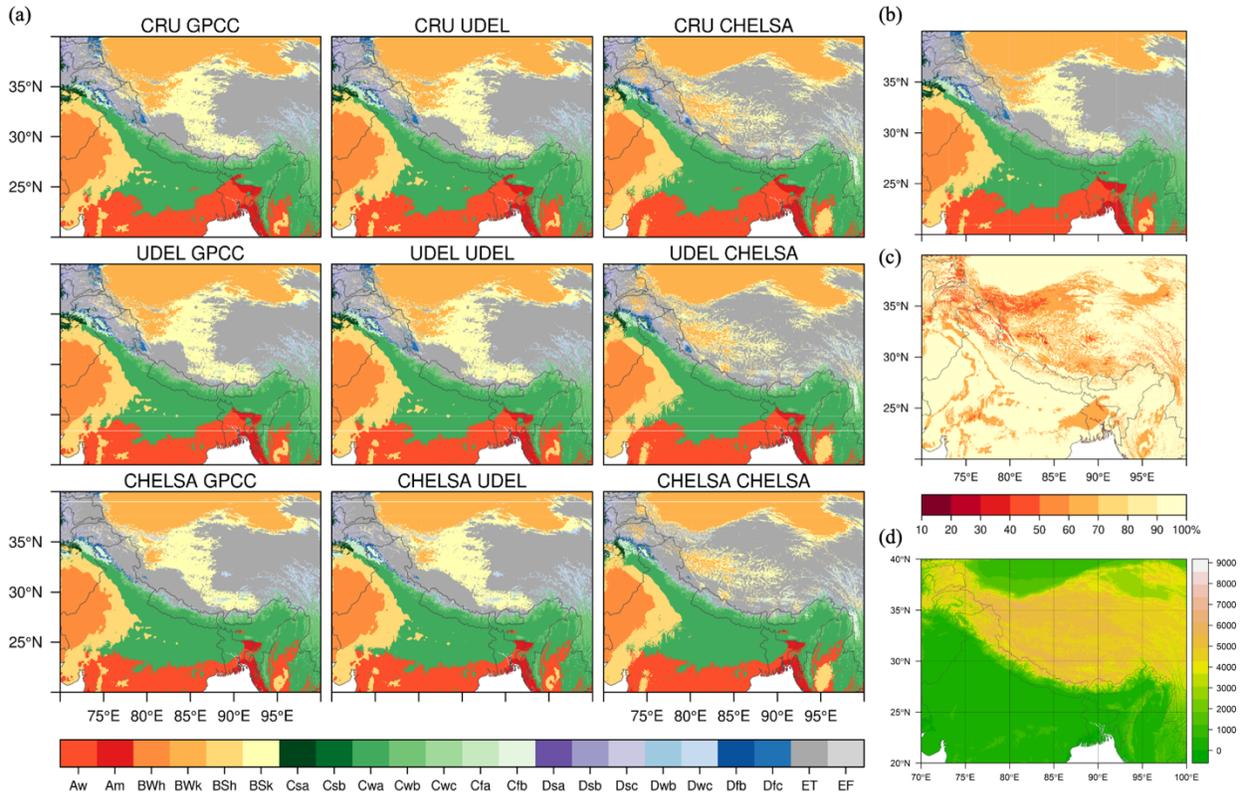
Line 224: maybe worth calling it an average climate classification map given the climate could have changed over this time period?

We totally agree. The climate classification maps describe the average climatic conditions over the given time period and are expected to change over a long time period. We have illustrated the point about the changes in KG climate in Introduction and Section 4.7.

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

The confidence level indicates the agreement level of the classification results based on different combinations of climatology datasets. Lower confidence level means larger discrepancies in the climate data, and poor reliability of the climate map. The regions with lower confidence levels are mostly climate transitional zones, which are more susceptible to changes in climatic conditions. This creates more uncertainty in identification of changes in KG climate zones. The study aims to improve the accuracy of climate classification maps to promote its usage in climate change studies.

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?

The dataset of bias-corrected downscaled CMIP5 model simulations from CCAFS, used in future KGC maps, presented data for four RCP scenarios and four 30-yr future time periods: 2030s (2020-2049), 2050s (2040-2069), 2060s (2050-2079) and 2070s (2070-2099). The four future time periods with a time interval ranging from 10 to 20 years, are expected to provide detailed information of the changing climate throughout this century. Considering the time and storage required to obtain the scenario-by-period downscaled output from model simulations, it is computationally expansive to process every 30-yr period in the future 80 years.

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.

We agree that it is hard to identify the changes in KG zones by only showing the future distributions of KG climate. We have added Figure 6c and 6d to show the projected changes in KG zones by the end of this century and the corresponding confidence levels of the areas which are projected to undergo changes.

Thanks for this suggestion. We have added Figs. 4 and 5 b in Figs. 7 and 8.

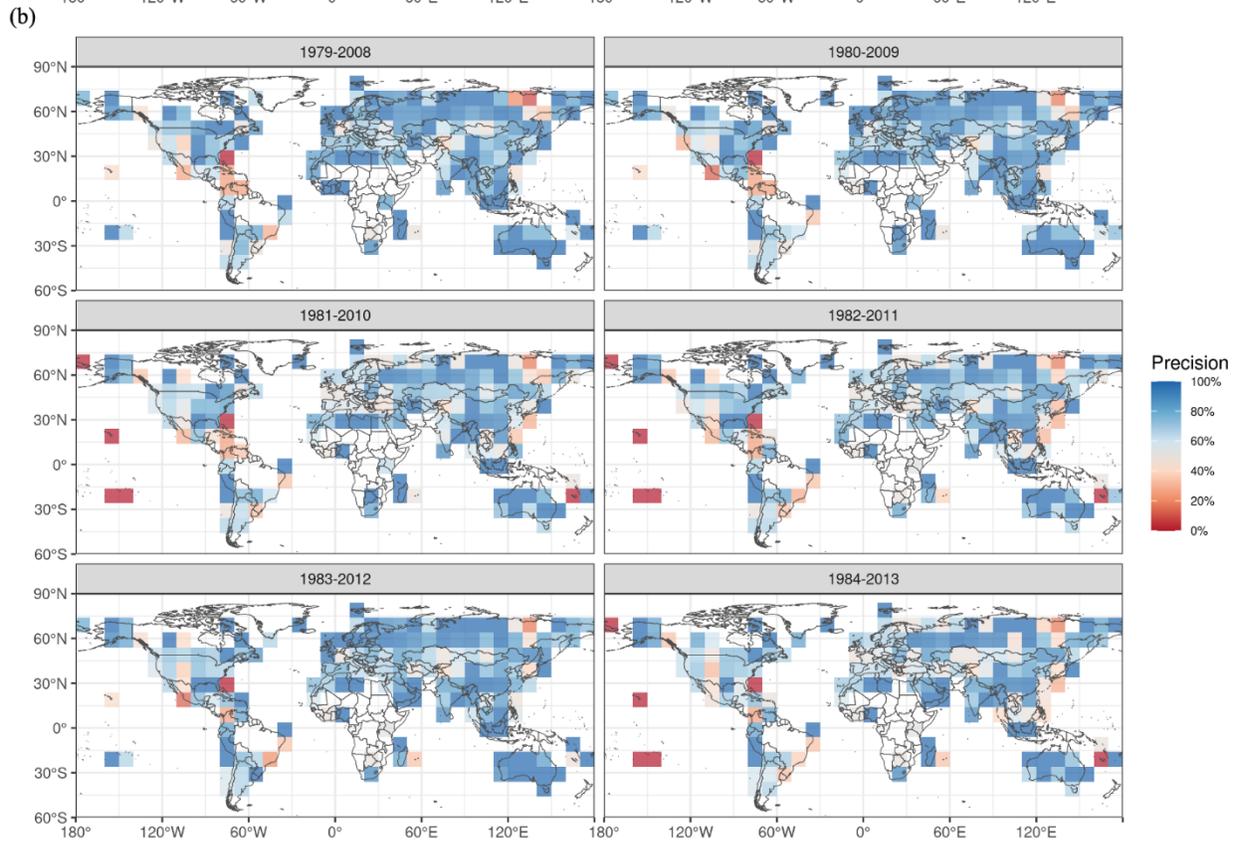
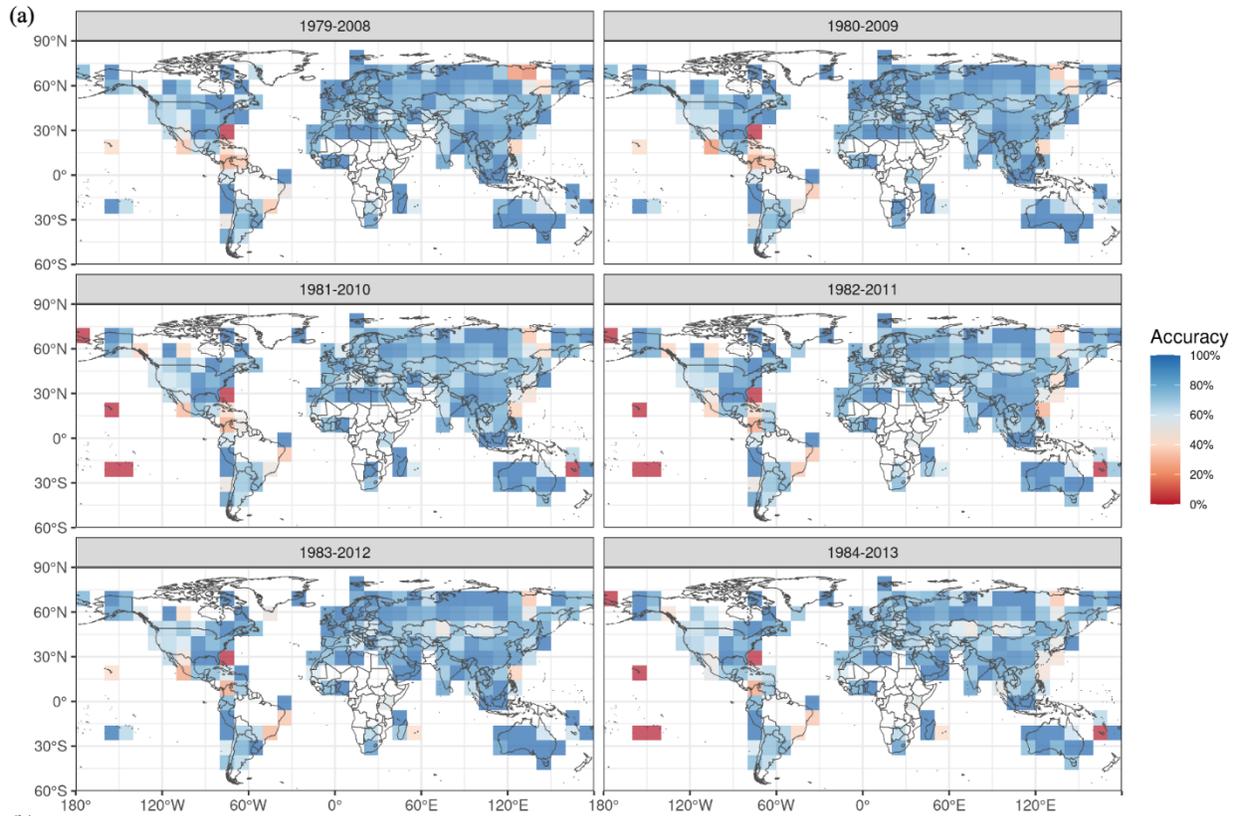
Section 4.3 and Figure 9: What is the overall accuracy and precision of the historical maps? I cannot see this reported anywhere, only the different combinations of datasets in Section 4.4. It might also be worth reporting accuracy and precision ranges of the historical map per region, or another more 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.

We added Table 3 in the revised manuscript, which shows the continental and global overall accuracy, average precision, and confidence level values of the historical maps.

Table 3 Continental and global overall accuracy, average precision, and confidence level of the historical Köppen-Geiger climate map series (1979-2008, 1980-2009, 1981-2010, 1982-2011, 1983-2012, 1984-2013). The overall accuracy is calculated as the percentage of correct climate classes using ground observations, and average precision is averaged fraction of correct classification for all climate classes. Confidence level values shows the 95% confidence interval of the confidence level for each continent, and the whole globe. All the values are presented in percentage.

Region	Africa	Asia	Oceania	Europe	North America	South America	Global	
Accuracy	1979-2008	88.24%	84.05%	92.39%	85.11%	79.37%	69.18%	83.25%
	1980-2009	87.67%	85.00%	90.11%	84.24%	76.94%	70.00%	82.96%
	1981-2010	85.71%	84.29%	93.48%	84.23%	75.61%	68.75%	82.63%
	1982-2011	83.78%	85.06%	91.30%	84.10%	74.79%	68.90%	82.42%
	1983-2012	85.43%	83.64%	92.39%	83.51%	71.99%	66.67%	81.48%
	1984-2013	85.81%	81.32%	92.39%	84.38%	71.84%	68.00%	81.62%
Precision	1979-2008	80.24%	72.77%	92.77%	75.71%	64.41%	66.20%	71.27%
	1980-2009	88.33%	73.40%	89.83%	75.58%	65.15%	68.11%	73.39%
	1981-2010	79.54%	71.19%	94.21%	74.77%	67.75%	67.63%	74.10%
	1982-2011	70.42%	71.34%	91.37%	75.61%	70.62%	66.65%	74.24%
	1983-2012	71.54%	68.99%	92.67%	69.82%	66.73%	64.33%	72.41%
	1984-2013	71.66%	68.08%	92.55%	76.30%	67.95%	65.17%	74.59%
Confidence level	1979-2008	94.93±0.002%	92.08±0.002%	91.82±0.002%	92.29±0.002%	94.55±0.004%	92.31±0.003%	92.94±0.002%
	1980-2009	94.91±0.002%	92.14±0.002%	91.73±0.002%	92.39±0.002%	94.65±0.004%	92.24±0.003%	92.95±0.002%
	1981-2010	94.89±0.002%	92.17±0.002%	91.63±0.002%	92.43±0.002%	94.51±0.004%	92.18±0.003%	92.92±0.002%
	1982-2011	94.92±0.002%	92.16±0.002%	91.48±0.002%	92.41±0.002%	94.35±0.004%	92.13±0.003%	92.87±0.002%
	1983-2012	94.96±0.002%	92.16±0.002%	91.31±0.002%	92.54±0.002%	94.37±0.004%	92.05±0.003%	92.87±0.002%
	1984-2013	94.97±0.002%	91.22±0.002%	91.32±0.002%	92.52±0.002%	94.45±0.004%	92.00±0.003%	92.87±0.002%

For Figure 9, we changed the color bars to make the colors more distinct, as shown below:

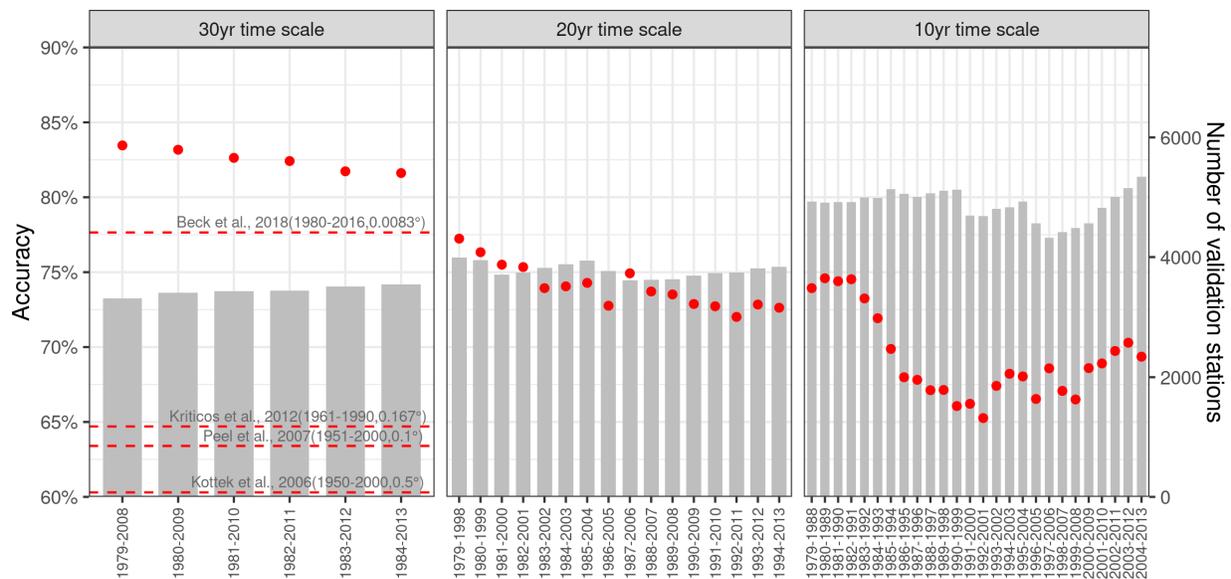


In Section 4.4 the total accuracy is still using the same validation datasets as in Section 4.3, am I correct? This needs to be a little clearer in the text, and 4.4 This needs to be made a little clearer in the text.

Correct. We have clarified this in section 4.4.

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.

Yes, the total accuracy of Beck et al., (2018), Kriticos et al., (2012), Peel et al., (2007), and Kottek et al., (2006) was calculated using the same validation datasets. We used the same classification system described in the previous studies and the same time period of the previous climate map product to process the station observation data and estimate the overall accuracy. We have added the explanation of the validation method in the revised manuscript. For the missing accuracy of Beck et al. (2018), we have added the line to represent the Beck et al. accuracy in the updated figure, as attached below:



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?

We accepted your suggestion and combined Sections 4.3 and 4.4 into one section, which is entitled “Validation”.

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).

Thanks for your advice. Previously we described the validation methods in Section 4 along with the validation results. We agree that it will be better if we move this part to methodology section and provide more details about the validation dataset selection, sensitivity analysis and so on. To address the needs, we added a new sub-section to Section 3, which is Section 3.4, to describe the methods used for technical validation, sensitivity analysis, and product comparison. Below is the added content of the new sub-section:

“We validated the historical climate maps using the station observations from Global Historical Climatology Network-Daily (GHCN-D) (Menne, Durre, Vose, Gleason, & Houston, 2012) and Global Summary Of the Day (GSOD) database (National Climatic Data Center, NESDIS, NOAA, & U.S. Department of Commerce, 2015) as reference data. GHCN-D dataset provides daily climate data over global land areas and contains records from over 80,000 weather stations worldwide, about one third of which have both temperature and precipitation data available (Menne et al., 2012). GSOD dataset includes global daily summary data over 9,000 stations, of which the historical data from 1973 being the most complete (National Climatic Data Center et al., 2015). For each station, time series of monthly temperature and precipitation were calculated from the daily observations with months with <15 daily values discarded. Then if ≥ 6 months are present, monthly climatology were generated subsequently by averaging the monthly means for the given 30-yr period. We removed duplicate stations in the two datasets and discarded stations with gap years or missing data in the given 30 years. For each station and each 30-yr period, we applied the Köppen-Geiger climate classification, and then evaluated overall classification performance for each climate map using total accuracy, which is defined as the percentage of correct classes, and average precision, which is averaged fraction of correct classification for all climate classes.

Using the same validation datasets and station selection process, we also evaluated the previous climate maps from Beck et al., (2018) Kriticos et al., (2012), Peel et al., (2007), and Kottek et al., (2006). We applied the same Köppen-Geiger climate classification criteria described in the previous studies to assess the overall accuracy of the map products. To further validate the climate classification results, we performed sensitivity analysis on the data integration method, the climate classification time scale, and climatology dataset input, using the same validation datasets from GHCN-D and GSOD. In addition, we compared the climate classification results with forest cover and elevation maps, and with the two high-resolution comparable climate map products, Beck et al., (2018) (1-km) and Kriticos et al., (2012) (0.167 °), at regional and continental scale.”

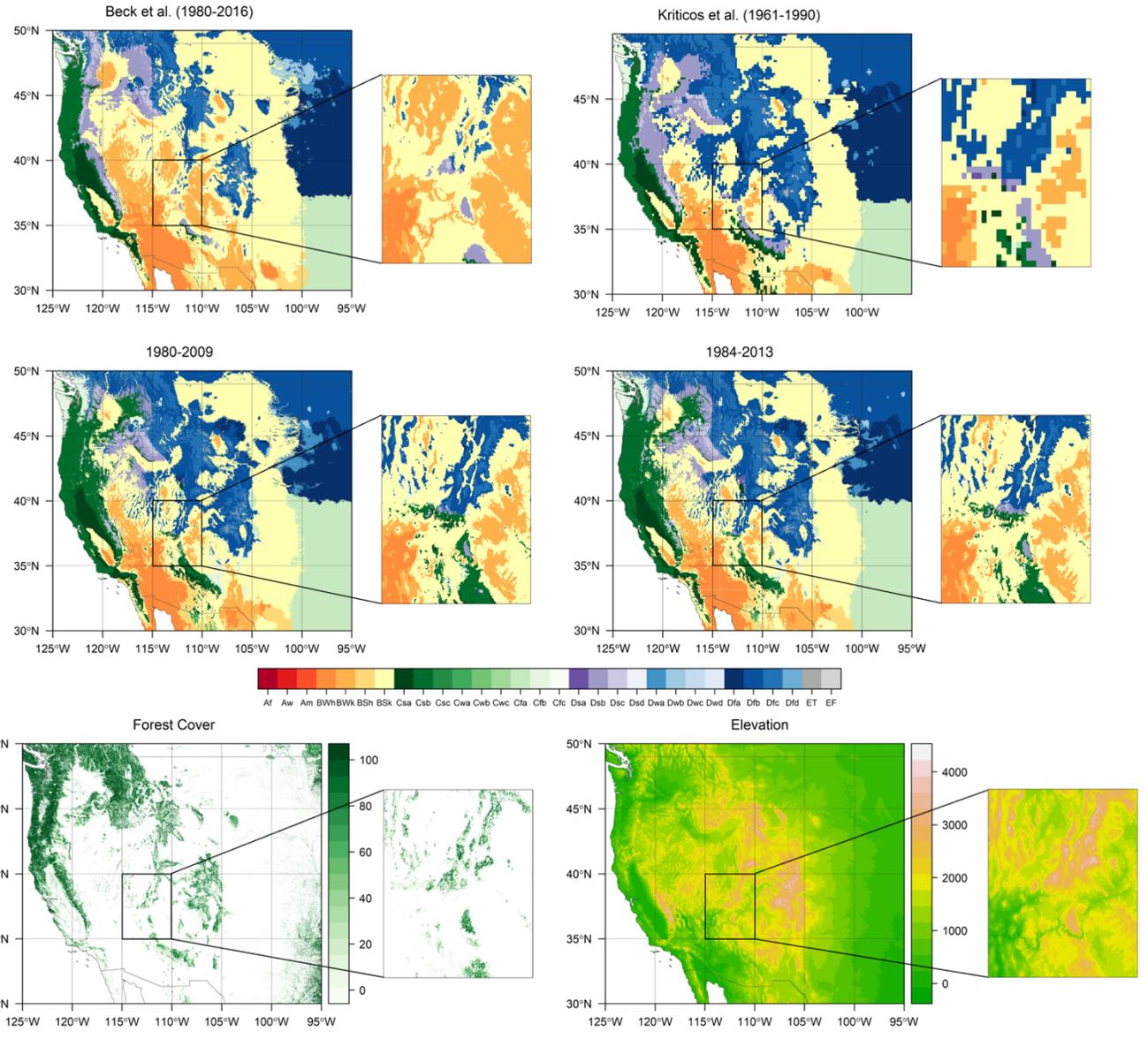
Also the methods used to derive Figure 10, including determining the accuracy of the other datasets, is not described in the methods. This is not clear and yet is important for highlighting how this product is better. Descriptions of confidence level, accuracy and precision calculations, sensitivity 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?

We added Section 3.4 to describe the validation, confidence level, accuracy and precision calculations, and sensitivity analysis. The details are included in the response to the previous comment. Thanks for pointing out the issue about the use of accuracy in the manuscript. We have revised it to confidence level when it indicates the confidence level value.

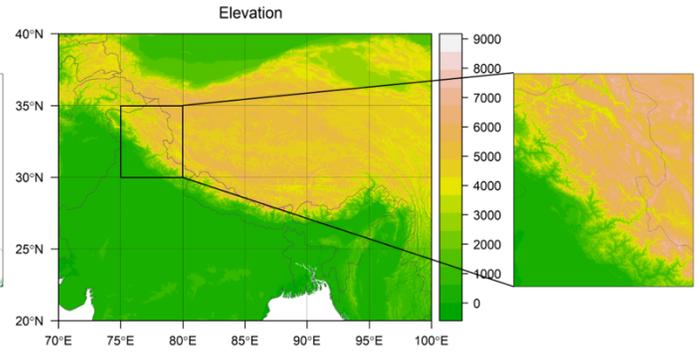
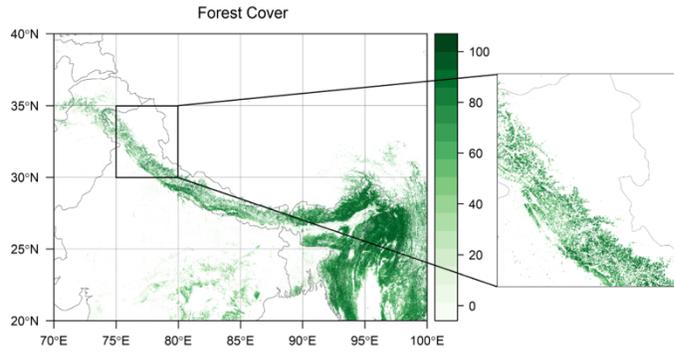
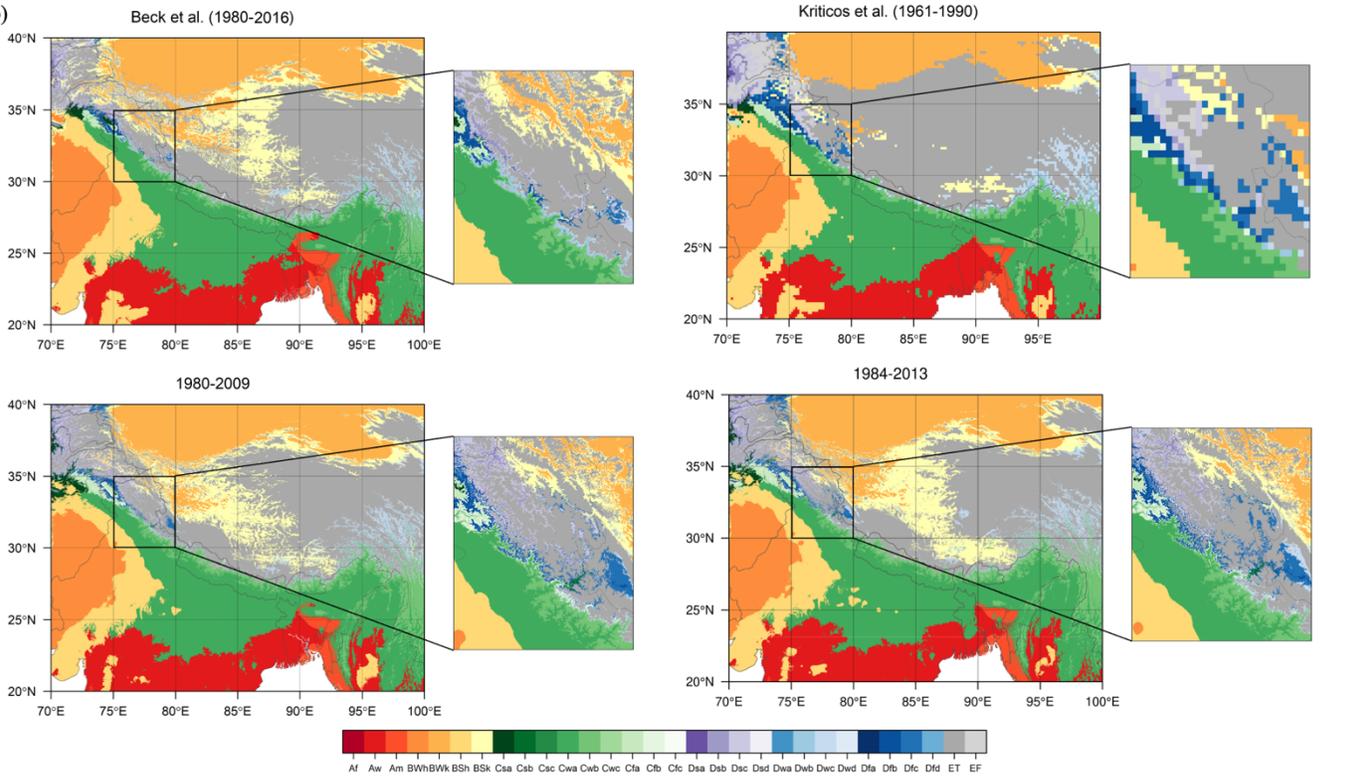
The differences between different KG maps, and how we can use forest cover to assess those maps, is not that clear to me in Figure 11 and the associated text. For example these two statements in line 310-311 and lines 318-319 is not backed up well by figure 11: “Figure 11 illustrate the enhanced regional 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.

Thanks for raising the concern about the KG map assessment at regional scale. To address the comment, first, we added the justification for the map assessment using forest cover map, and we updated Figure 11 to show the continental and regional scale comparison with more details. We also included more careful description of the comparison in the Figure. Here are the updated figure and the revised version of the section:

(a)

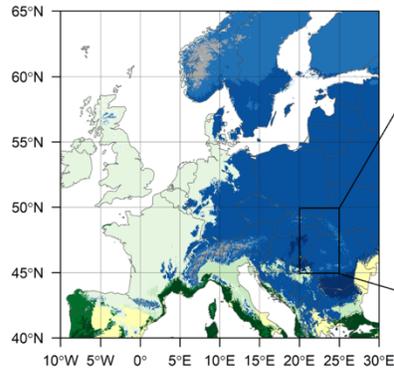


(b)

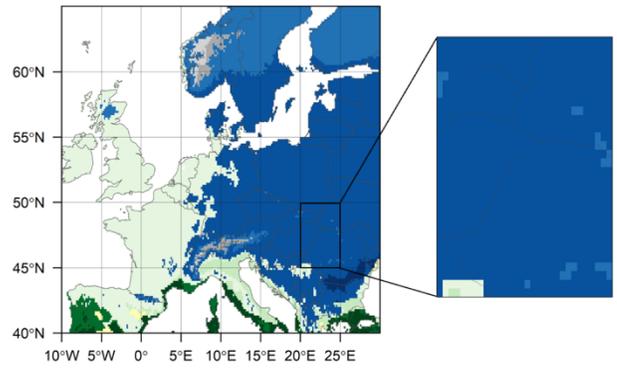


(c)

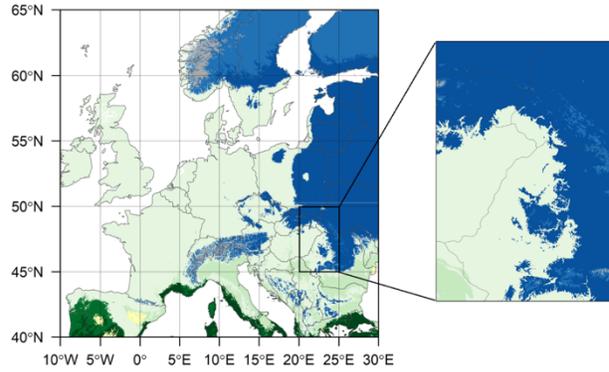
Beck et al. (1980-2016)



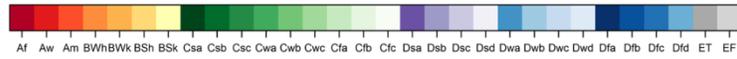
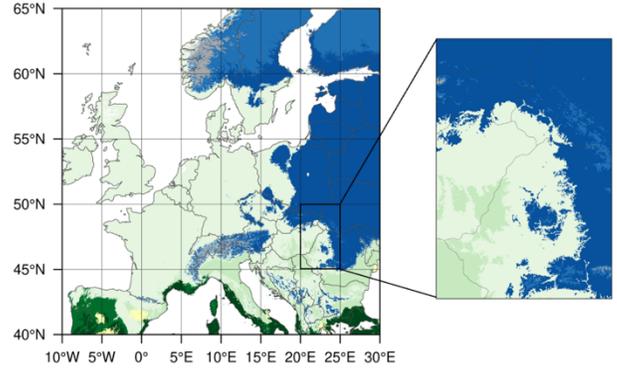
Kriticos et al. (1961-1990)



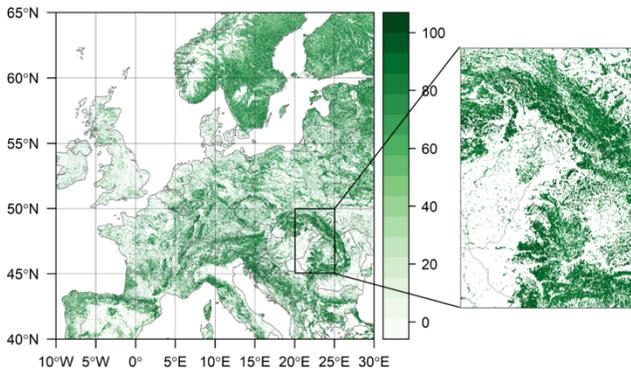
1980-2009



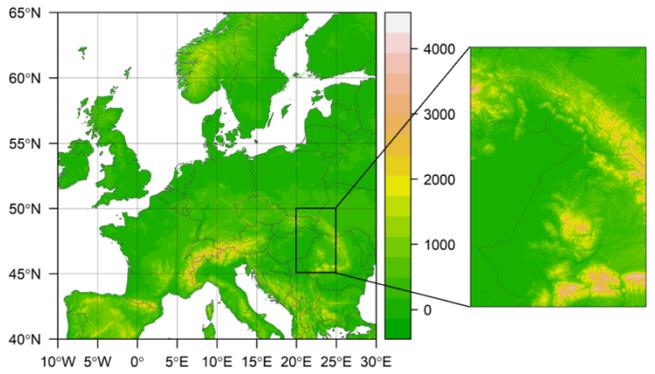
1984-2013



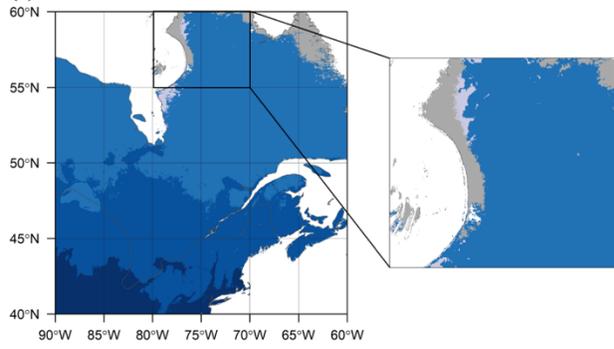
Forest Cover



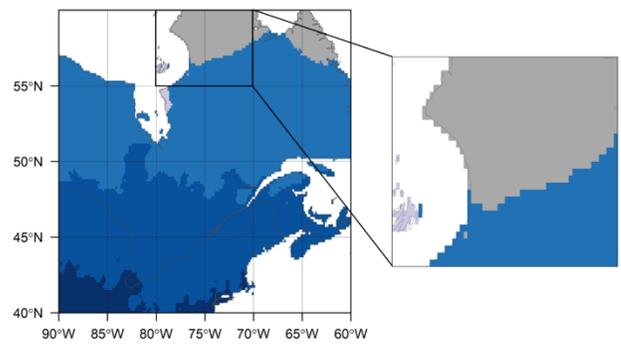
Elevation



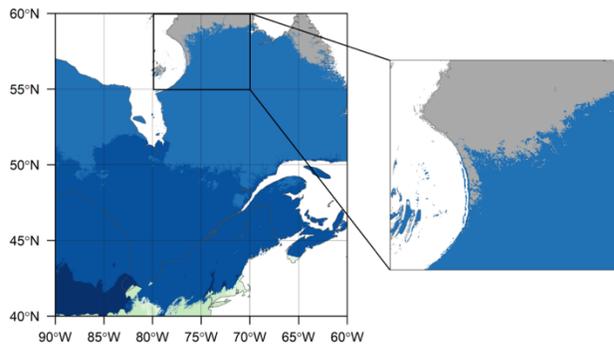
(d) Beck et al. (1980-2016)



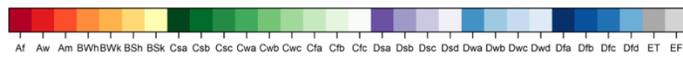
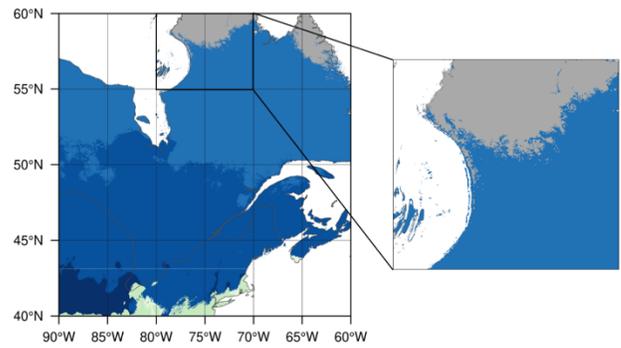
Kriticos et al. (1961-1990)



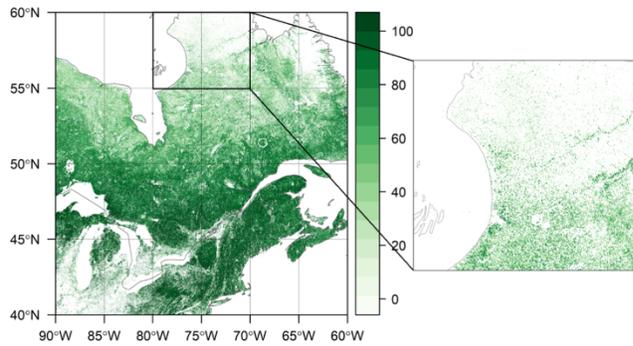
1980-2009



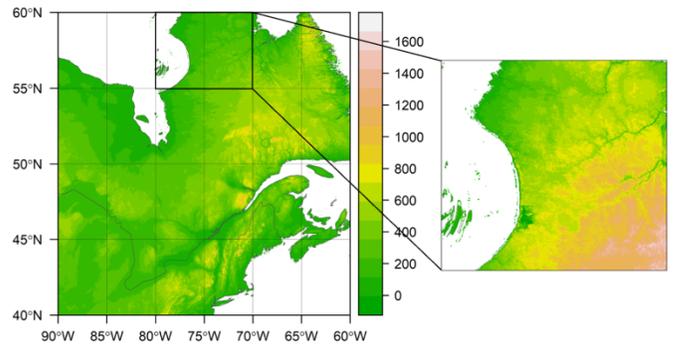
1984-2013

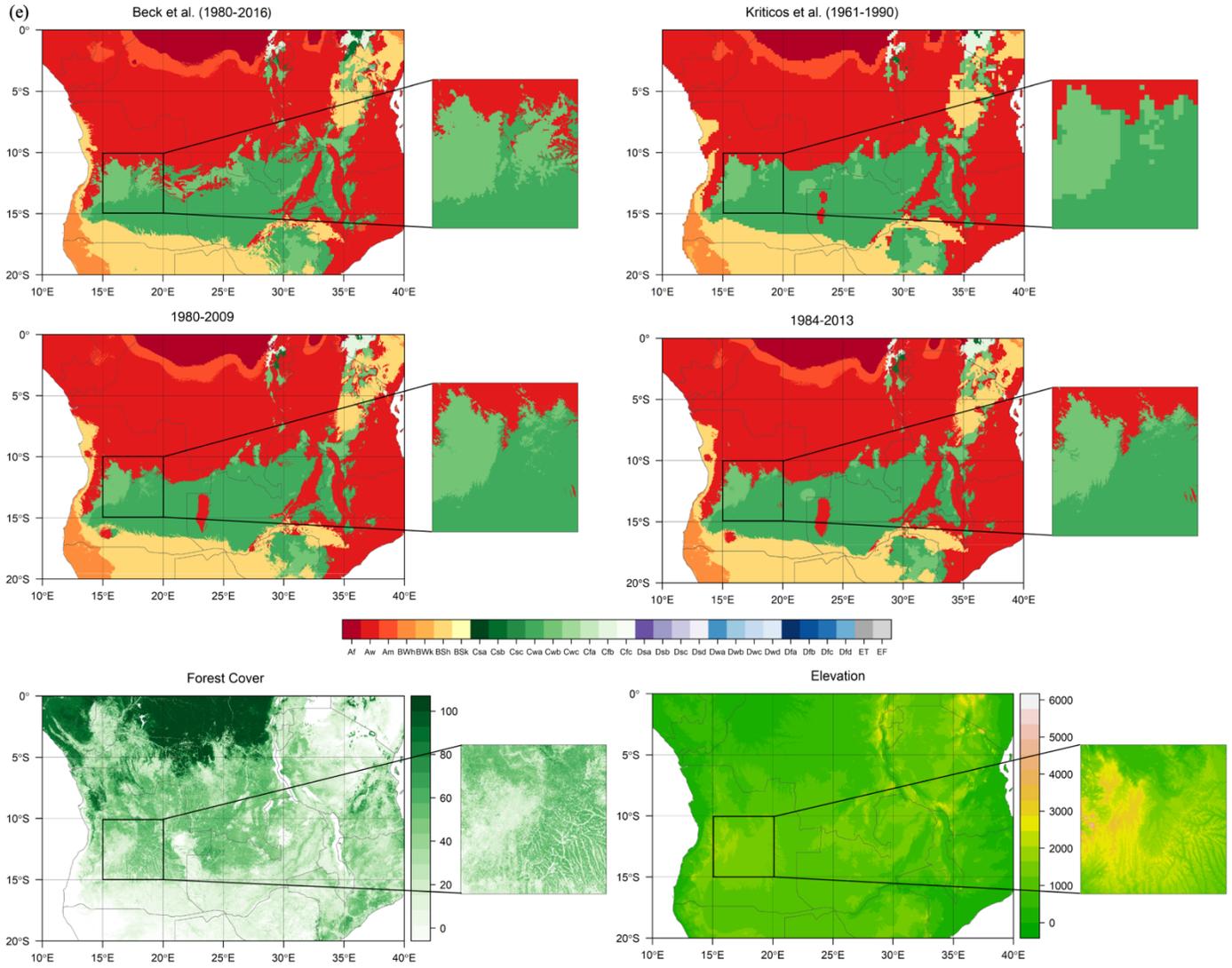


Forest Cover



Elevation





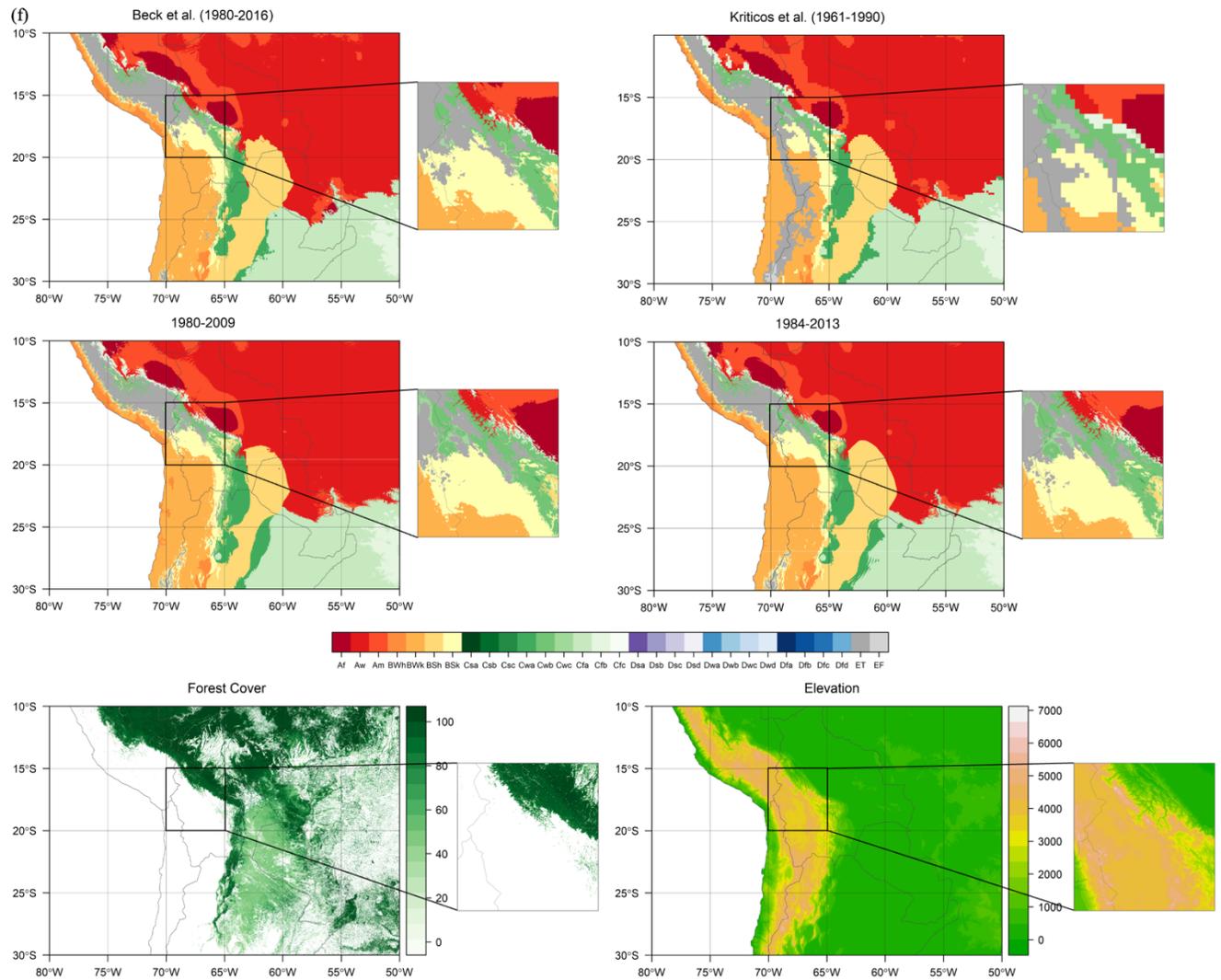


Figure 8. Köppen-Geiger climate classification maps from previous studies, Beck et al., 2018 (1-km, 1980-2016), and Kriticos et al., 2012 (0.167°, 1961-1990), our study (1-km, 1979-2009 to 1984-2013), associated forest cover and elevation maps, for regions with large spatial gradients in climates or sharp elevation gradients. (a) central Rocky Mountains, (b) Tibetan Plateau, (c) Europe, (d) high latitudes in North America, (e) Central and eastern Africa, and (f) central Andes. The forest cover map is the 30m Landsat-based forest cover map for year 2000 (Hansen et al., 2013). The elevation data is the NASA SRTM Digital Elevation 30m data (Farr et al., 2007). The representative period of each map is listed in parentheses.

“At the regional and continental scale, we compared our Köppen-Geiger climate classification maps with previous map products for regions with large spatial gradients in climates, including central and eastern Africa, Europe, North America, and regions with sharp elevation gradients, including Tibetan Plateau, central Rocky Mountains, central Andes (Fig. 8). We compared the new 1-km Köppen-Geiger climate classification maps from our study for time periods of 1980-2009, and 1984-2013 with the high-resolution Köppen-Geiger maps from two previous studies, Beck et al., (2018), which has a resolution of 1-km and temporal coverage of 1980-2016, and Kriticos et al., (2012), which has a resolution of 0.0167° and covers 1961-1990. The Köppen classifications demonstrate good correlation with natural landscape distributions (Belda et al., 2014; Köppen, 1936; Trewartha, 1954). To show the agreement between the improved Köppen-Geiger climate classification maps and

regional landscape distributions, we also showed maps of forest cover, and elevation distribution for these regions. Figure 8 illustrate the enhanced regional details of the maps.

Compared with the Köppen-Geiger climate maps from previous studies with only one time period, the series of the Köppen-Geiger climate maps from our study demonstrate the ability to capture recent changes in spatial distributions of climate zones. For example, our maps can detect the significant changes in the climate zones specifically driven by the accelerated global warming since the 1980s, for example, the poleward movements of boreal (D) and polar (E) climates in high latitudes in North America shown in the comparison between the 1980-2009 and 1984-2013 Köppen-Geiger climate maps (Fig. 8d). Another example is the expansion of savanna (Aw) climate into temperature (Cw) climate zone, witnessed in Central Africa (Fig. 8e).

Another improvement of the new series of the Köppen-Geiger climate maps is the application of threshold of -3 oC as the boundary of temperate (C) and boreal (D) climate zones, which show better agreement with global boreal forest distributions at regional scale compared with Russell's modification of 0 oC (1931), which Beck et al., (2018), and Kriticos et al., (2012) utilized. Based on the comparison results of the Köppen climate zones and the biome classifications from the World Wildlife Federation (Rohli et al., 2015), the boreal (D) climate zone largely corresponds to the distribution of boreal forest (Cui, Liang, & Wang, 2021). For example, evidenced in Figure 8c, the new Köppen-Geiger climate classification maps from our study show better agreement with the boreal forest in Carpathian Mountains across Central and Eastern Europe than Beck et al., (2018), and Kriticos et al., (2012). Figure 8d also shows good agreement of the northern boundary of boreal (D) climate zone in northern part of Quebec in Canada with the boundary of Canada's boreal forest.

Moreover, the new Köppen-Geiger maps can show accurate depiction of important topographic features over the regions with complex topography. For example, the topoclimate of the Himalays southern front determined by the mountain ranges are represented with more details in the new Köppen-Geiger maps compared with Beck et al., (2018), and Kriticos et al., (2012) (Fig. 8b). The abrupt changes in climate along the edges of the Andes mountains are also well described in the new maps (Fig. 8f).

In addition, the distribution of tropical (A), temperate (C) and boreal(D) climate zones in the new Köppen-Geiger maps correspond closely with tree lines in the forest cover maps. The temperate (C) and boreal(D) climate distributions based on the Köppen-Geiger maps show a better agreement with the forest distributions of the Middle and Southern Rocky Mountains than Beck et al., (2018), and Kriticos et al., (2012) (Fig. 8a). For another example, the boundaries of the tropical rainforest in Central Africa and South America are clearly delineated in the in the new Köppen-Geiger maps (Fig. 8e and 8f)."

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.

Thanks for this valuable suggestion. We agree that the reorganization provides a better structure for section 4. We have moved sections 4.3-4.6, before 4.2 and 4.7 in the revised manuscript.

Technical Corrections

A number of “the”s missing throughout the text.

Thanks for pointing it out. We added a number of “the”s in the manuscript and will pay closer attention to the use of it.

References

- Beck, C., Grieser, J., & Rudolf, B. (2005). A new monthly precipitation climatology for the global land areas for the period 1951 to 2000. *Klimastatusbericht KSB, 2004*.
- Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data, 5*, 180214. <https://doi.org/10.1038/sdata.2018.214>
- Belda, M., Holtanová, E., Halenka, T., & Kalvová, J. (2014). Climate classification revisited: From Köppen to Trewartha. *Climate Research, 59*(1), 1–13. <https://doi.org/10.3354/cr01204>
- Cui, D., Liang, S., & Wang, D. (2021). Observed and projected changes in global climate zones based on Köppen climate classification. *Wiley Interdisciplinary Reviews: Climate Change*, e701. <https://doi.org/10.1002/wcc.701>
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., . . . Alsdorf, D. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics, 45*(2), 1485. <https://doi.org/10.1029/2005RG000183>
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology, 37*(12), 4302–4315. <https://doi.org/10.1002/joc.5086>
- Funk, C., Verdin, A., Michaelsen, J., Peterson, P., Pedreros, D., & Husak, G. (2015). *A global satellite assisted precipitation climatology*.
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., . . . Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. *Science (New York, N.Y.), 342*(6160), 850–853. <https://doi.org/10.1126/science.1244693>
- Hay, L. E., Wilby, R. L., & Leavesley, G. H. (2000). A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the united states. *JAWRA Journal of the American Water Resources Association, 36*(2), 387–397. <https://doi.org/10.1111/j.1752-1688.2000.tb04276.x>
- Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., . . . Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data, 4*, 170122. <https://doi.org/10.1038/sdata.2017.122>
- Köppen, W. P. (1936). *Das geographische System der Klimate: Mit 14 Textfiguren*: Borntraeger.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift, 15*(3), 259–263. <https://doi.org/10.1127/0941-2948/2006/0130>

- Kriticos, D. J., Webber, B. L., Leriche, A., Ota, N., Macadam, I., Bathols, J., & Scott, J. K. (2012). CliMond: Global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods in Ecology and Evolution*, 3(1), 53–64. <https://doi.org/10.1111/j.2041-210X.2011.00134.x>
- Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., & Houston, T. G. (2012). An overview of the global historical climatology network-daily database. Advance online publication. <https://doi.org/10.1175/jtech-d-11-00103.1>
- National Climatic Data Center, NESDIS, NOAA, & U.S. Department of Commerce (2015). Global Surface Summary of the Day - GSOD. Retrieved from <https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.ncdc:C00516#>
- Navarro-Racines, C., Tarapues, J., Thornton, P., Jarvis, A., & Ramirez-Villegas, J. (2020). High-resolution and bias-corrected CMIP5 projections for climate change impact assessments. *Scientific Data*, 7(1), 7. <https://doi.org/10.1038/s41597-019-0343-8>
- New, M., Hulme, M., & Jones, P. (2000). Representing twentieth-century space–time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate. *Journal of Climate*, 13(13), 2217–2238. [https://doi.org/10.1175/1520-0442\(2000\)013<2217:RTCSTC>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2)
- Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen–Geiger climate classification. *Hydrology and Earth System Sciences Discussions*, 4(2), 439–473. <https://doi.org/10.5194/hessd-4-439-2007>
- Rohli, R. V., Joyner, T. A., Reynolds, S. J., & Ballinger, T. J. (2015). Overlap of global Köppen–Geiger climates, biomes, and soil orders. *Physical Geography*, 36(2), 158–175. <https://doi.org/10.1080/02723646.2015.1016384>
- Rubel, F., & Kottek, M. (2010). Observed and projected climate shifts 1901–2100 depicted by world maps of the Köppen–Geiger climate classification. *Meteorologische Zeitschrift*, 19(2), 135–141. <https://doi.org/10.1127/0941-2948/2010/0430>
- Russell, R. J. (1931). *Dry climates of the United States: I. Climatic map* (Vol. 5): University of California press.
- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K.-L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, 56(1), 79–107. <https://doi.org/10.1002/2017RG000574>
- Trewartha, G. T. (1954). *An introduction to climate*: McGRAW-HILL BOOK COMPANY, INC. NEW YORK TORONTO LONDON.
- Wilby, R. L., & Wigley, T.M.L. (1997). Downscaling general circulation model output: A review of methods and limitations. *Progress in Physical Geography: Earth and Environment*, 21(4), 530–548. <https://doi.org/10.1177/030913339702100403>
- Willmott, C. J., & Matsuura, K. (2001). Terrestrial air temperature and precipitation: monthly and annual time series (1950 - 1999). Retrieved from http://climate.geog.udel.edu/~climate/html_pages/README.gncn_ts2.html.