The authors used GLASS CDRs data and Google Earth Engine platform and produced the long-term continuous land cover dataset from 1982 to 2015. This is a very valuable dataset for further applications in the analyses of energy and carbon dynamics and the global land surface modelling. However, I have some concerns about the data processing in the classification, accuracy assessment and the interpretation of the land cover change results. I think these problems must be solved / addressed before publication.

We thank the reviewer for the comments and thoughtful review. Please find our detailed response along with the suggested changes to our manuscript below.

1. Differences between forest and tree cover

The authors used vegetation cover fraction (VCF) data from Song et al. 2018. However, in their paper, they specified "tree cover" increase. This is not equal to forest increase. Usually, the forest is defined by canopy closure (e.g. tree cover fraction >10% in FAO, >25% in Hansen et al. 2013), tree height and minimum area. The authors showed that a lot of forest increase occurred in Siberia (Fig. 10) and was from grassland (Fig. 11). This could be artificial considering the coarse resolution (5 km) and poor ability of land cover mapping for mosaic pixels (see below) in the methods used in this study. For example, there is 5 ha forest with tree cover fraction of 35%, and the tree cover fraction increased to 45% because of better growth (e.g. longer growing season, CO_2) in the same 5 ha forest. In this case, we cannot say the forest area increased 5 ha x 10% = 0.5 ha because it is the same 5 ha forest but with denser tree cover. Therefore, I doubt that there is confusion of these concepts in this manuscript and maybe in the classification system. The authors briefly mentioned this issue on L450-454, but this really needs to be clarified, assessed and solved.

Response 1:

Thank you for your advice. It should be pointed out that our classification target is land cover class, not vegetation cover percentage information. Our land cover products belong to the hard classification and give each mapping unit a single land cover class. VCF is only used as features that assist in the land cover classification, which is introduced as prior probability, and only one of many factors that affect the final classification result. Although based on VCF information, our results are not the same.

The classification system we used is from FROM-GLC_v2 (Li et al., 2017). Considering the quality of the data and the separability of classes, our products include 7 land cover classes, cropland, forest, grassland, shrubland, tundra, barren land, and snow and ice (Table 1). Among them, the forest is also defined and distinguished by canopy closure. The forest is defined under the condition that tree cover≥10% and height> 5m. We have updated the description of the classification system in our manuscript.

Table 1: Classification system, with 7 Level 1 classes and 21 Level 2 classes.

Level 1 class	Level 2 class	Description
Cropland	Rice paddy Greenhouse Other farmland Orchard Bare farmland	
Forest	Broadleaf, leaf-on Broadleaf, leaf-off Needle-leaf, leaf-on Needle-leaf, leaf-off Mixed leaf type, leaf-on Mixed leaf type, leaf-off	Tree cover≥10%; Height>5m; For mixed leaf, neither coniferous nor broadleaf types exceed 60%
Grassland	Pasture, leaf-on Natural grassland, leaf-on Grassland, leaf-off	Canopy cover≥20%
Shrubland	Shrub cover, leaf-on Shrub cover, leaf-off	Canopy cover≥20%; Height<5m
Tundra	Shrub and brush tundra Herbaceous tundra	
Barren land	Barren land	Vegetation cover<10%
Snow/Ice	Snow Ice	

We agree with you that forest increase may exist under the condition that you described. This is an inevitable problem in hard classification. What we call forest increase is the change of land cover class in our classification results under our 5km coarse resolution classification system. Limited to a spatial resolution of 5km, there are many mixed mapping units. For these mixed units, the estimation of hard classification will cause a large deviation. This is a common problem in hard classification. Similar problems also exist in land cover data prediction with higher spatial resolution. At coarse resolution, accurate estimates may be better with cover percentage data.

Change in manuscript:

We have updated the description to our used classification system in Table 1.

2. The majority land cover in a 0.05 deg pixel

The majority method in a coarse resolution (5 km) may work for some pure pixels but is expected to work poorly for the mosaic pixels with high heterogeneity or similar fraction of different vegetation types. For example, in a 5 km pixel with 43% tree cover, 44% grass and 13% others in the first year, it became 45% tree cover, 44% grass and 11% of others in the second year simply because of the good climate. If I understood correctly, this pixel would be classified as grassland in the first year and forest in the second year, and thus there is a 25 km2 land cover change from grassland to forest. This may also partly explain the strong forest increase in Siberia, high variations in the temporal land cover dynamics in Fig. 8 and the high uncertainties in the intensive LCC regions (e.g. savanna in Africa).

Response 2:

Thanks for your comment. For mosaic pixels, especially mosaic pixels of vegetation, hard classification does have such disadvantages. The classification system used in the MODIS-based land cover product has included some mosaic classes, such as the Forest / Cropland Mosaics, Natural Herbaceous / Croplands Mosaics and Herbaceous Croplands defined in the FAO-Land Cover Classification System land use (LCCS2) system, which also reflects the difficulty and disadvantage of hard classification in coarse resolution to a certain extent. However, for these mosaic classes in the MODIS-based land cover product, hard classification is still used. Although at individual pixel level this is unavoidable when land cover data are aggregated over large areas the extreme cases as raised by the reviewer would usually be averaged out.

Despite of the disadvantage, the way that hard classification presents information is more direct. In many applications, researchers prefer to use the results of hard classification.

Besides, the scheme we used to aggregate and extract coarse-resolution samples from fine-resolution data is one of the common used schemes (DeFries et al., 1998;Wang et al., 2016). Under the framework of hard classification, there does not seem to be a better solution.

As for the LCC area reflected in the product, there are some places, as you said, that may be affected by the hard classification method. However, there are also many areas where the LCC is correctly reflected, such as the forest area of the Amazon region cut back. It should be pointed out that the LCC information in our results has uncertainty, especially the regions with high variability in LCC.

Change in manuscript:

We have added a reminder to data users about the uncertainty of our products.

3. Accuracy of change detection

The authors only assessed the accuracy for year 2015, not mentioning that the uncertainty of FROM-GLC_v2 was not propagated. First, the same product was used for training the classification system and for the accuracy assessment. Although the samples in the same product may be not overlapped, we cannot exclude the coherence since both are from FROM-GLC_v2. So, some independent evaluation dataset would be helpful. Second, an important feature of this continuous land cover maps is the temporal dynamics. So, the change detection needs to be further validated / evaluated in addition to the one-year classification accuracy assessment. This part is currently lacking in this work.

Response 3:

Thank you for your useful comment. In this revision, we collected a new independent test sample and performed the accuracy assessment. To prove the impact of change detection, we further compared the accuracies with and without change detection.

Specifically, we collected 2431 randomly distributed 5km sample points in different years around the world. According to the majority principle, we manually interpreted the land cover class of each sample as an independent test sample. Besides, to verify the accuracy of the change detection method, we also compared the classification accuracy before and after the change detection. The temporal distribution of the newly collected test samples is shown in Fig. 1, and the geographical distribution is shown in Fig. 2.

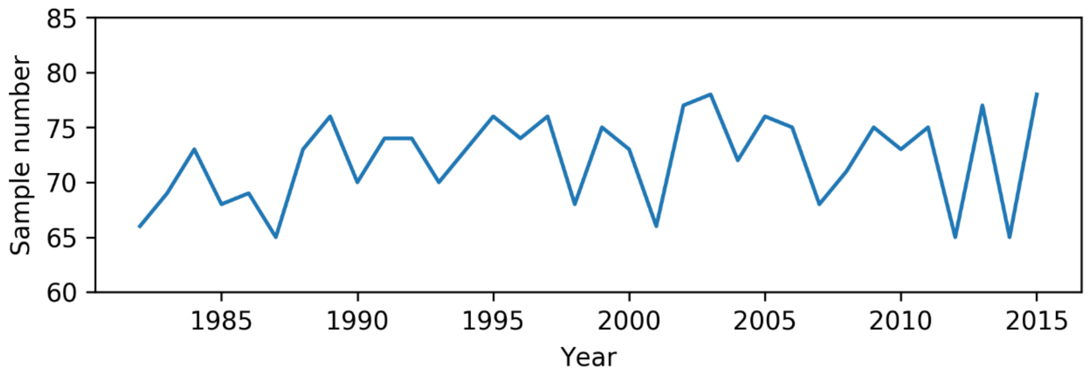

Figure 1: The temporal distribution of the newly collected test sample.

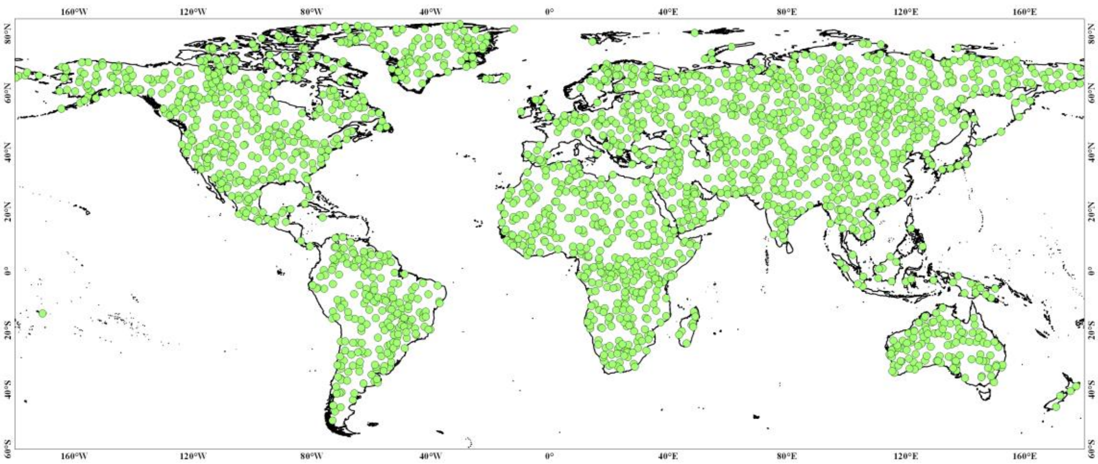

Figure 2: The geographical distribution of random test sample.

The new assessment result is shown in Table 3 and Table 4. It shows that OA of GLASS-GLC without change detection is 81.28%, and OA with change detection is 82.81%. This reflects the reliability of GLASS-GLC since the test samples are randomly distributed along the spatial and temporal dimensions, and also confirm the significance and effectiveness of the change detection method.

Table 3: Classification accuracy of GLASS-GLC without change detection under 2431 independent test samples. (Overall accuracy = 81.28 %, UA = User's Accuracy and PA = Producer's Accuracy)

Class	Cropland	Forest	Grassland	Shrubland	Tundra	Barren land	Snow/ice	Total number	UA
Cropland	257	21	34	15	0	31	0	358	71.79%
Forest	35	620	45	27	22	1	1	751	82.56%
Grassland	17	26	248	12	3	19	4	329	75.38%
Shrubland	7	6	10	154	9	12	0	198	77.78%
Tundra	0	9	11	12	250	3	0	285	87.72%
Barren land	4	1	13	14	5	355	6	398	89.20%
Snow/ice	0	4	3	0	0	13	92	112	82.14%
Total number	320	687	364	234	289	434	103	2431	
PA	80.31%	90.25%	68.13%	65.81%	86.51%	81.80%	89.32%		81.28%

Table 4: Classification accuracy of GLASS-GLC with change detection under 2431 independent test samples. (Overall accuracy =82.81 %, UA = User's Accuracy and PA = Producer's Accuracy)

Class	Cropland	Forest	Grassland	Shrubland	Tundra	Barren land	Snow/ice	Total number	UA
Cropland	262	19	32	20	0	25	0	358	73.18%
Forest	33	637	29	28	24	0	0	751	84.82%
Grassland	24	24	254	6	13	8	0	329	77.20%
Shrubland	12	3	11	159	6	7	0	198	80.30%
Tundra	0	12	9	4	250	10	0	285	87.72%
Barren land	5	1	17	8	7	357	3	398	89.70%
Snow/ice	0	5	6	0	0	7	94	112	83.93%
Total number	336	701	358	225	300	414	97	2431	
PA	77.98%	90.87%	70.95%	70.67%	83.33%	86.23%	96.91%		82.81%

Change in manuscript:

We have added the new accuracy assessment result in the manuscript.

4. Comparison with other datasets

A suggestion for the evaluation may be to compare the total area, spatial and temporal changes with other datasets e.g. ESA-CCI 300 m, Hansen forest, FAO and some cropland datasets. This would help to verify the mapping results in this study and to understand their differences. It would also help to define the possible applications of this dataset (e.g. whether it can be used for carbon accounting, land modeling).

Response 4:

Thank you for your advice. Comparison with other land cover products is a very good way to reflect product quality and accuracy. For this reason, in addition to the classification accuracy obtained by several evaluation methods, we compared other available land cover products with our products. Although there are some differences in the classification system of different products, it can still reflect the reliability of our products in general.

We inter-compared GLASS-GLC with other available global land cover products with a relatively long time series. Land cover products from MODIS and the ESA-CCI were used. The MODIS-based global land cover products come from Collection 6 (C6) MODIS Land Cover Type (MLCT) products (Sulla-Menashe et al., 2019), and are supervised classification results from 2001 to 2016. Considering the comparability to our classification system, the FAO-Land Cover Classification System land use (LCCS2) layer was used. The corresponding relationships of classes are listed as follows, and the class names we used are the latter: barren - barren land, permanent snow and ice – snow/ice, all kinds of forest – forest, forest/cropland mosaics and natural herbaceous/cropland mosaic – cropland, natural herbaceous and herbaceous cropland – grassland, shrubland - shrubland. The ESA-CCI global land cover products (Bontemps et al., 2013) are 300m resolution yearly products ranging from 1992 to 2015. The products were developed using the GlobCover unsupervised classification chain and merging multiple available Earth observation products based on the GlobCover products of the ESA (Liu et al., 2018). Referring to the class relationships in (Liu et al., 2018), we cross-walked classes including cropland, forest, grassland, shrubland, barren land and snow/ice.

Apart from land cover products, we also compared GLASS-GLC with the Food and Agricultural Organization of the United Nations statistical data (FAOSTAT) on cropland and forest (forest land) classes, which are the main sources of country-level land cover data for many applications. The annual FAOSTAT data set on cropland we used ranged from 1982 to 2015, and that on forest we used ranged from 1990 to 2015.

We made an inter-comparison between classes including cropland, forest, grassland, shrubland, barren land and snow/ice. The main inter-comparison is the area corresponding to the top 50 countries in each class. Besides, to compare the accuracy of different products, test samples from FLUXNET site data in 2015 are given for independent accuracy assessment.

The assessment results of MODIS-based land cover products and ESA-CCI land cover products based on test samples from FLUXNET site data are shown in Table 6 and Table 7, respectively. The overall accuracies of ESA-CCI products and MODIS-based products are 73.90% and 80.38% in 2015, respectively. Compared to these, The overall accuracy of GLASS-GLC (82.10%, Table 5) is superior. Although the cross-walk of the different classification systems may be slightly different, It can still reflect the high accuracy of our GLASS-GLC products.

Table 5: Classification accuracy of GLASS-GLC in 2015 based on FLUXNET test sample. (Overall accuracy = 82.10 %, UA = User’s Accuracy and PA = Producer’s Accuracy)

Class	Cropland	Forest	Grassland	Shrubland	Tundra	Barren land	Snow/ice	Total number	UA
Cropland	**63**	5	17	1	0	0	0	86	73.26 %
Forest	13	**243**	9	2	0	0	0	267	91.01 %
Grassland	8	21	**91**	2	0	2	0	124	73.39 %
Shrubland	7	3	0	**19**	0	0	0	29	65.52 %
Tundra	0	3	0	0	**14**	0	0	17	82.35 %
Barren land	0	1	0	0	0	**1**	0	2	50.00 %
Snow/ice	0	0	0	0	0	0	**0**	0	-
Total number	91	276	117	24	14	3	0	**525**	
PA	69.23 %	88.04 %	77.78 %	79.17 %	100.00 %	33.33 %	-		**82.10 %**

Table 6: Classification accuracy of the MODIS-based land cover product in 2015 based on FLUXNET test sample. (Overall accuracy = 82.10 %, UA = User's Accuracy and PA = Producer's Accuracy)

Class	Cropland	Forest	Grassland	Shrubland	Tundra	Barren land	Snow/ice	Total number	UA
Cropland	**7**	5	73	0	0	0	0	85	8.24%
Forest	1	**261**	5	0	0	0	0	267	97.75%
Grassland	1	15	**108**	1	0	0	0	125	86.40%
Shrubland	0	9	9	**11**	0	0	0	29	37.93%
Tundra	0	3	6	8	**0**	0	0	17	-
Barren land	0	0	1	0	0	**1**	0	2	50.00%
Snow/ice	0	0	0	0	0	0	**0**	0	-
Total number	9	293	202	20	0	1	0	**525**	
PA	77.78%	89.08%	53.47%	55.00%	-	100.00%	-		**73.90%**

Table 7: Classification accuracy of the ESA-CCI land cover product in 2015 based on FLUXNET test sample. (Overall accuracy = 82.10 %, UA = User's Accuracy and PA = Producer's Accuracy)

Class	Cropland	Forest	Grassland	Shrubland	Tundra	Barren land	Snow/ice	Total number	UA
Cropland	**81**	1	4	0	0	0	0	86	94.19%
Forest	11	**246**	4	5	0	1	0	267	92.13%
Grassland	28	7	**76**	5	0	8	0	124	61.29%
Shrubland	2	7	1	**19**	0	0	0	29	65.52%
Tundra	0	3	9	0	**0**	5	0	17	-
Barren land	0	0	2	0	0	**0**	0	2	0.00%
Snow/ice	0	0	0	0	0	0	**0**	0	-
Total number	122	264	96	29	0	14	0	**525**	
PA	66.39%	93.18%	79.17%	65.52%	-	0.00%	-		**80.38%**

Figure 3 shows an inter-comparison with MODIS-based products, Figure 4 with ESA-CCI products and Figure 5 with FAOSTAT. The scatter plots and the linear fit lines reflect the results in 2015, and the box plots represent the distribution of R^2 of the annual linear fit lines for each class. It can be seen that various classes in several different products are relatively equivalent although they are under different classification systems. In comparison with MODIS-based products, the results of 2001-2015 for cropland, forest and snow/ice have high R^2. In comparison with ESA-CCI products, the mean R^2 of the linear fit lines of forest, grassland and snow/ice from 1992 to 2015 reach 0.99, 0.82, and 0.98, respectively, while the R^2 for shrubland is low. The inter-comparison of some other classes is poor, which may be caused by differences in the class definition in various classification systems. For instance, our classification system incorporates tundra, while the other two did not. Compared with FAOSTAT, the mean R^2 of the linear fit lines of cropland and forest is 0.82, and 0.87, respectively. In general, our GLASS-GLC products have a reasonable consistency with other products and statistics and the difference are not significant.

What's more, the duration of GLASS-GLC is much longer than MODIS-based and ESA-CC land cover products (as shown in Fig. 6). The comparison with other data illustrates the reliability and superiority of GLASS-GLC.

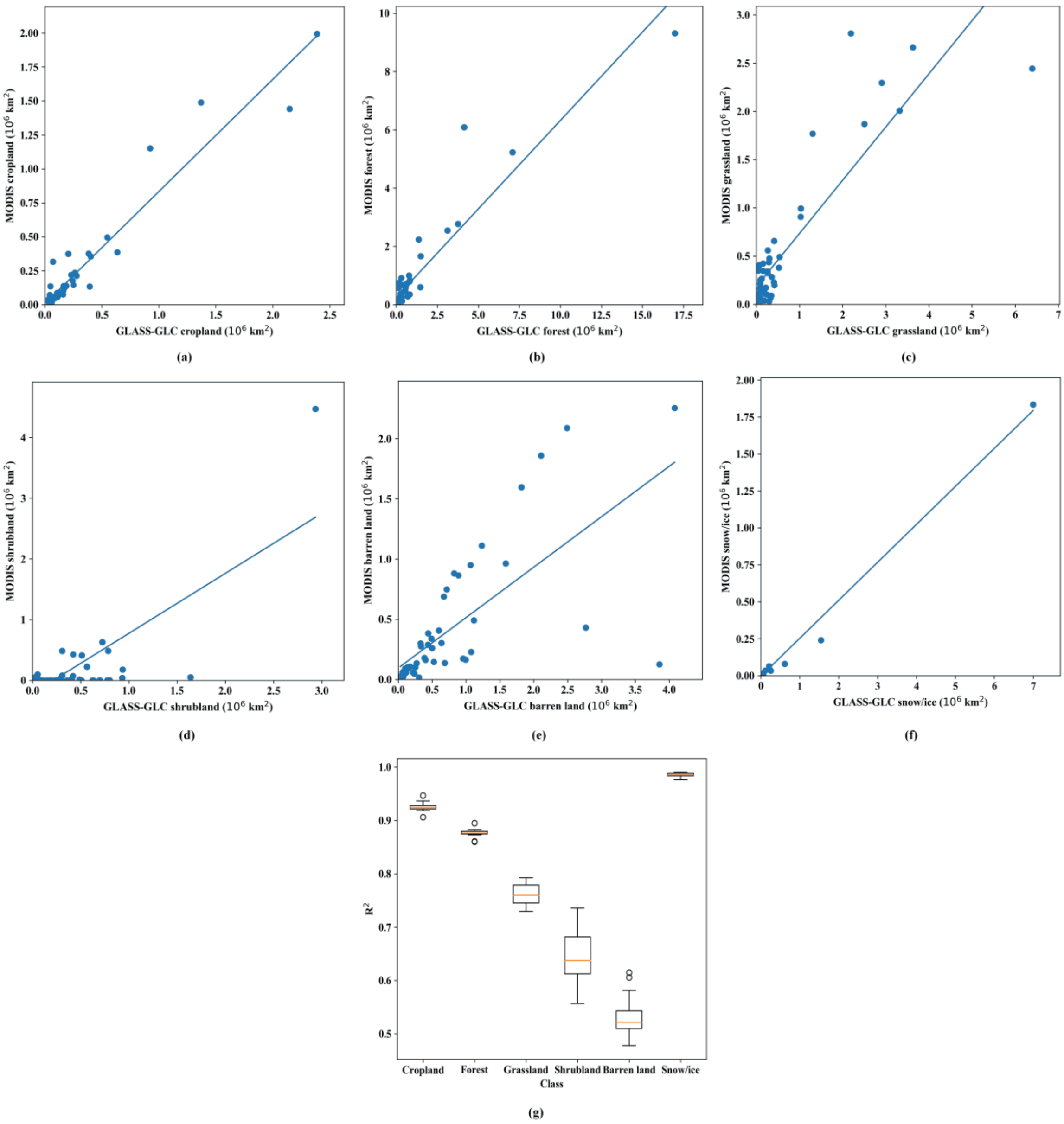

Figure 3: Inter-comparison with the MODIS-based land cover product, (a) cropland circa 2015, (b) forest circa 2015, (c) grassland circa 2015, (d) shrubland circa 2015, (e) barren land circa 2015 and (f) snow/ice circa 2015; (g) mean R2 of the annual linear fit lines for all years (2001-2015).

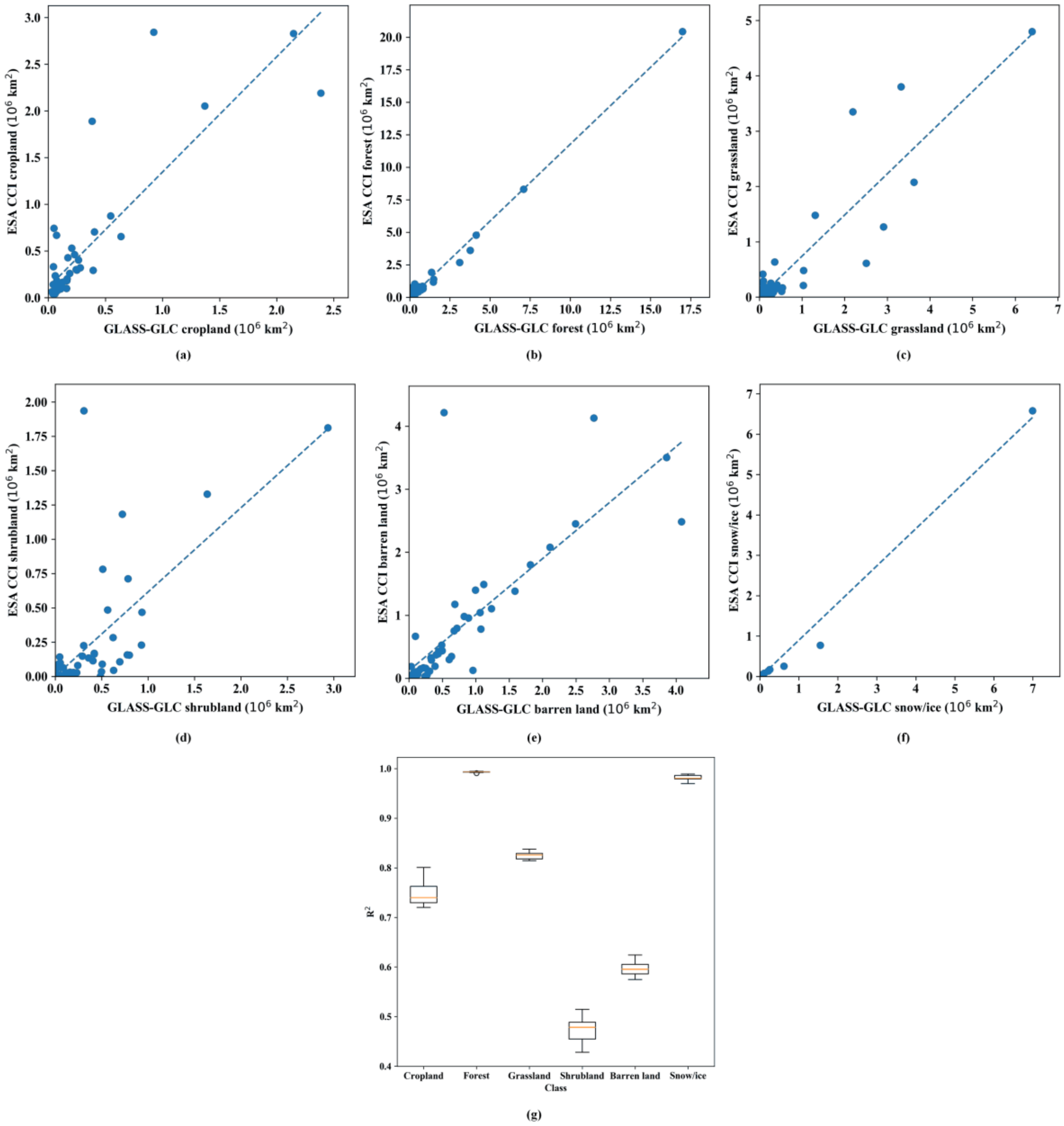

Figure 4: Inter-comparison with the ESA-CCI land cover product, (a) cropland circa 2015, (b) forest circa 2015, (c) grassland circa 2015, (d) shrubland circa 2015, (e) barren land circa 2015 and (f) snow/ice circa 2015; (g) mean R2 of the annual linear fit lines for all years (1992-2015).

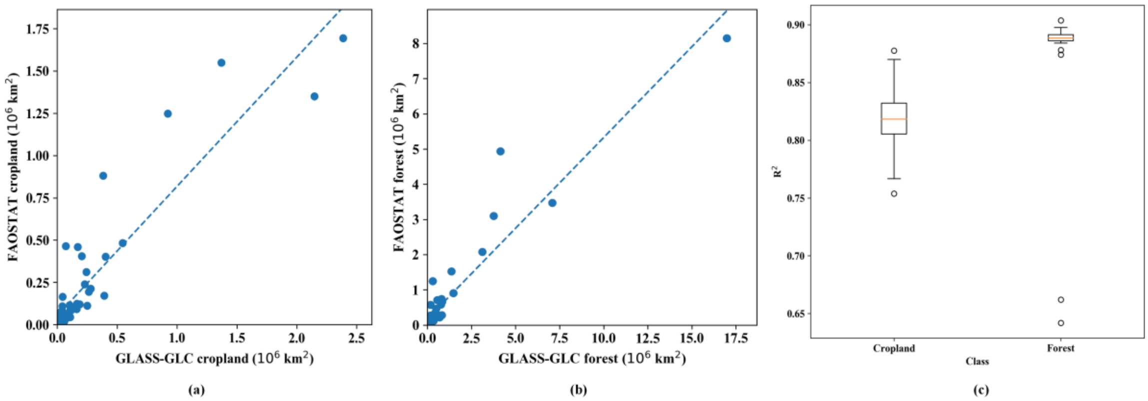

Figure 5: Inter-comparison with the FAOSTAT data set, (a) cropland circa 2015, (b) forest circa 2015, (c) mean R2 of the annual linear fit lines of cropland for years 1982-2015 and forest for years 1990-2015.

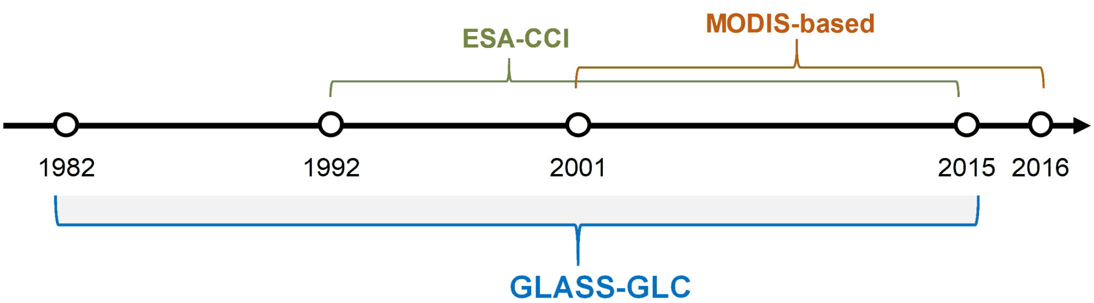

Figure 6: The duration of different land cover products, including GLASS-GLC, MODIS-based land cover products and ESA-CCI land cover products.

Change in manuscript:

We have included comparison results with other land cover data in the manuscript to help show the reliability and effectiveness of our products.

5. Superficial and fragmented interpretations of reasons for LCC

The authors made a lot of figures and tables to show the spatial and temporal changes and also reasons for such changes. These sections are not well organized and lack some internal logics. What I learned is only some fragmented information. The reasons for the LCC are not very solid (see my detailed comments below). Just taking one example, the author mentioned several times of "greening" and its effects on LCC. However, greening is very far away from LCC. It may only be caused by more leaves and extended growing season. We don't know whether this increased productivity was converted to carbon stock or leaded to a land cover transition from grass to forest. The increased carbon uptake by greening may just release back to the atmosphere through the enhanced respiration due to increased temperature. So, I would suggest being cautious when interpreting the reasons for the LCC. In fact, I don't think these sections are necessary for this manuscript. Adding comparisons with other datasets and discussing the differences between various data and the reasons (e.g. data sources, classification methods) would be enough for a nice data paper. The reasons for LCC can be separated to another paper after adding more analyses. Putting it here only attenuate the main objective of this manuscript.

Response 5:

Thank you for your comment. The interpretations of reasons for LCC are just some examples of our attempts to apply our product for further analysis, not the main focus of this paper. The focus of this article is on the presentation and quality assessment of our produced GLASS-GLC data products. To this end, we have added more content on accuracy assessment and product inter-comparison, to better demonstrate the reliability and uncertainty of our products. As for the reasons for LCC, we will analyze and discuss in more detail in a subsequent paper.

Change in manuscript:

We have supplemented the sections of accuracy assessment and data inter-comparison.

6. Writing
Language needs further improvements. A lot of sentences are difficult to understand, and some sentences are broken in the context. Please polish the language during revision.

Response 6:

Thanks for your suggestion.

Change in manuscript:

We have polished our language with a native English consultant.

Specific comments:
L19: report how many classes

Response s1:

The classification system consists of 7 classes, including cropland, forest, grassland, shrubland, tundra, barren land, snow/ice, as shown in Table 1 in the manuscript.

Change in manuscript:

We have added the information in the corresponding place as you suggested.

L20: 85% accuracy based on what?

Response s2:

It was based on 23459 test samples in 2015. And the overall accuracy of the produced GLASS-GLC CDR in 2015 is 86.51 %. The test samples come from the 30 m resolution FROM-GLC_v2 test sample set (Li et al., 2017).

To give a more effective assessment, we also performed an accuracy assessment using FLUXNET site data and the newly collected independent test samples, and we supplemented this part of the results.

Change in manuscript:

We have updated the detail in the corresponding place.

L22: how can you separate afforestation and forest expansion?

Response s3:

The data products we produce can only provide information at the observation level. For example, the information we can obtain here is only forest gain. While the specific causes of these LCCs should be analyzed and investigated separately, we cannot distinguish afforestation and natural expansion of forests.

Change in manuscript:

We have modified our description according to our study.

L23: land degradation? did you mean grassland loss? if it is degradation, it may still be grassland.

Response s4:

Yes, we do mean by grassland loss. At the individual mapping unit level, we cannot detect land degradation.

Change in manuscript:

We have changed the word “land degradation” to “grassland loss”.

L25: greening is not directly related to LCC. very complex processes behind.

Response s5:

Greening is indeed a very complex process. Here, we mainly refer to vegetation gain such as forest gain in our results, which can only be used as evidence from the perspective of remote sensing and mapping. Thanks for pointing it out.

Change in manuscript:

We have corrected our expression.

L37: What is “surface attributes”?

Response s6:

It refers to the characteristics and properties of the Earth surface. The change of land cover would change the status of the Earth surface.

Change in manuscript:

We have changed the word “attributes” to “characteristics”.

L44: too strong statement

Response s7:

Thank you for your reminding.

Change in manuscript:

We have modified our expression.

L70: “which will…” useless half sentence

Response s8:

Many thanks.

Change in manuscript:

We have rechecked the sentence and deleted it.

L75: not clear “more prone to consistency and data volume”, rephrase

Response s9:

What we mean here was that Landsat data has a higher spatial resolution, but it also meets some problems including obvious cloud contamination, data inconsistency caused by multiple generations of sensors and relatively larger data volume because of the high resolution.

Change in manuscript:

We have changed it to detailed description.

L87: “Because of…” duplicate

Response s10:

Thanks for your comment.

Change in manuscript:

We have changed the used “Because of” to “Due to”.

L91: analyses

Response s11:

Thank you for correcting us.

Change in manuscript:

We have changed the word.

L136: explain if you have level 2 class and how they were derived

Response s12:

There are no Level 2 classes in our results. Considering the resolution and separability of GLASS data, only Level 1 classes are included. The description of Level 2 classes comes from the original design in the FROM-GLC classification system (Gong et al., 2013). It was listed in Table 1 to better show the meaning of each Level 1 class. In the future, we will try to produce land cover products with a more detailed classification system.

L142: “2013-2015” is it a one-year map or three maps each for a year?

Response s13:

It was a one-year map, not three maps. Due to the problem of data quality, the Landsat data in one year usually cannot meet the need for land cover mapping on a global scale. The production of the FROM-GLC_v2 map took advantage of data from 2013 to 2015. And it can roughly be called circa 2015 (Li et al., 2017).

L145: “with a limited …” not clear what it is.

Response s14:

When generating random points in ArcGIS with the “create random points” tool, we limited the spatial interval among points greater than 0.1° by setting the parameter “minimum allowed distance” as 0.1.

Change in manuscript:

We changed our description.

L147: “class distribution” do you mean “percentage of each class”?

Response s15:

Yes. It means the percentage of each class.

Change in manuscript:

We have changed the description.

L158: what is end-number? end of what?

Response s16:

It is called end-member. Due to the complexity of ground objects and the limited spatial resolution of various sensors, the information contained in a pixel of remote sensing images is the mixture of information of many ground objects, hence resulting in mixed pixels (Zhang et al., 2011). It is assumed that there are pure land cover types known as basic mixing elements (known as end-member) that cannot be further decomposed in the imaged area, and the process of finding these end-members is referred to endmember extraction (Plaza et al., 2002). Here, to reduce the systematic deviation of AVHRR products, we correct the GLASS data with MODIS products based on end-members (Song et al., 2018).

L162: Is the smaller fluctuation the truth? Something you expected?

Response s17:

It is a trade-off. The purpose of data correction was to correct the original remotely sensed data to a higher consistency, especially in the temporal dimension. Remotely sensed data is easy to be affected by many random and systematic factors such as the atmospheric environment and sensor situation. The values it reflects were usually not those of the real and direct surface conditions. What's more, many fake inter-annual variations exist in remotely sensed data (Friedl et al., 2010). This will cause much trouble and disturbance especially in the use of time-series remotely sensed data. Though the variations may be caused by phenological changes, other interfering factors exist, and the trade-off is more beneficial in general.

The correction process carried out belongs to one of the data pre-processing processes in time-series land cover mapping (Gómez et al., 2016), which was to mitigate and deal with these aspects and to produce more consistent data for use.

L172: How is the performance of you trained random forest classifier? OOB R2 or independent evaluation dataset?

Response s18:

The OOB accuracy of our random forest classifier reached to 87.12%.

Change in manuscript:

We have added this information in the manuscript.

L174: what are the other parameters and the default values?

Response s19:

The specific parameters are listed as follows. The number of trees was 200, the out-of-bag mode is on. The number of variables per split was set to 0, as the square root of the number of variables. The minimum size of a terminal node was 1, the fraction of input to bag per tree was 0.5, and the random seed was 0.

Table 8 Specific parameters of the random forest classifier

Parameter	**Value**
Number of trees	200
Number of variables per split	0
Minimum size of a terminal node	1
Fraction of input to bag per tree	0.5
Whether the classifier should run in out-of-bag mode	True
Random seed	0

Change in manuscript:

We have listed the above parameter values in the manuscript.

L179: "the mode of…" not clear

Response s20:

The mode here refers to the class label that has the highest frequency in the segmented period with the calculated breakpoints. To improve the time consistency in the classification results, we use the mode class label to replace all the class labels in the period.

Change in manuscript:

We have updated our expression.

L186: How about the heterogeneous pixels? Not assessed at all?

Response s21:

Thanks for your question. In the newly collected independent test sample set, we use random points, with no difference between homogeneous and heterogeneous pixels. Therefore, the new assessment results include heterogeneous pixels.

L190: "class distribution"

Response s22:

Thanks for your reminding.

Change in manuscript:

We have changed the description.

L206: It is OK to fit a linear trend, but you cannot say to remove … because it may be caused by the actual LCC

Response s23:

Thanks for your advice. Our purpose was to fit a linear trend for better extraction of the land cover change trend in the long time-series land cover data. The fluctuations in the land cover were generally seen as an abnormal condition caused by climate conditions and phenological changes since the land cover is stable in most areas in the world across years, but they can be caused by the actual land cover change as you said.

Change in manuscript:

We have changed the description.

L212: why is summed?

Response s24:

Because we wanted to ensure the significance of the land cover change trend. For each pixel in the land cover map of each class, the original 0.05 ° pixel is labeled with 0 or 1 (belonging to the class or not), and such categorical data (not continuous data) cannot be statistically hypothesized. In order to carry out the hypothesis test, some studies used downscaling (Wang et al., 2016). By downscaling, the categorical label data can be summed up as the area ratio of the class (numerical data) in a greater statistical area, thus a statistical hypothesis test can be performed to verify the significance of land cover change.

L213: what is "annual change in slope of area ration"?

Response s25:

We are sorry for making a slip in writing. It is in fact "annual change slope of area ratio" estimated from a Theil-Sen estimator. More specifically, it represents the speed of land cover change.

Change in manuscript:

We have corrected it.

L219: why only statistically significant change was included? It is still area change even the trend is not significant. The way you process data exaggerate the changes.

Response s26:

Yes, there are certain shortcomings in doing so. But relatively speaking, this is a better strategy. Because it usually exists fake inter-annual land cover change in time-series land cover mapping studies (Sulla-Menashe et al., 2019) caused by many kinds of factors as explained in the above. Although there may be some real land cover change, to ensure the significance and reduce the uncertainty we did not include those into the statistics.

L223-224: Again, why only change mask?

Response s27:

Because we want to ensure the statistical significance and reduce the uncertainty caused by classification noises to detect more robust long-term land cover change trends.

L225: “direct” duplicate

Response s28:

Thank you for your comment.

Change in manuscript:

We have deleted the word.

L242: Need to explain UA and PA for non-remote-sensing readers; explain what the column and row names refer to.

Response s29:

Thanks for reminding. UA and PA represent user’s accuracy and producer’s accuracy respectively. They are two metrics reflecting the accuracy of classification. UA = corrected classified sample number / total sample number in the classification, PA = corrected classified sample number / total sample number in test sample.

Change in manuscript:

We explained the abbreviations in the titles of the corresponding tables.

L245: “Grassland is …” , from Table 3, they are shrubland and forest

Response s30:

It was concluded from the row dimension with the user’s accuracy. But as for the producer’s accuracy, it is as what you said.

L248: samples

Response s31:

Many thanks.

Change in manuscript:

We have corrected the word.

L255: these are regions with intensive LCC

Response s32:

Some regions such as Africa show relatively intensive LCC. There may be more mosaic pixels in these places in Africa, which may also lead to high uncertainty

For other regions with relatively high uncertainty, their locations are close to the continent edge which may be one of the reasons. The uncertainty map was reported based on the interpolation of test samples, the uncertainty values near the edges where test samples are rarely distributed would be affected to some degree.

L259: “variation curves” temporal changes

Response s33:

Thank you very much for your kindness.

Change in manuscript:

We have changed the phrase.

L260: Why so strong forest increase from 2006-2008? is it real?

Response s34:

We think it should be carefully treated. Since we do not have sufficient reference data, we cannot be sure if this is real or artifacts. The fluctuations in the curves can be seen as one of the representations of the uncertainty using coarse-resolution remotely sensed data.

L262: what about cropland? why so high variations, especially in 1994, 1999?

Response s35:

Cropland showed a slightly increasing trend, but not significantly. The high variations also reflect some kind of uncertainty introduced by the input data.

L263: Fig. 9, explain the meaning of your boxplot, mean, median, IQR, 90%, max, min? Why use the ratio, instead of total change area which is more straightforward?

Response s36:

The box extends from the first (lower) quartile (Q1) to third (upper) quartile (Q3) values of the data, with a line indicating the median. The whiskers extend from the box to show the range of the data. The upper whisker extends to the last datum less than Q3 + 1.5 * IQR, and the lower whisker extends to the first datum greater than Q1 - 1.5 * IQR. Flier points are those past the end of the whiskers. We wanted to use the change ratio to better reflect that how much percentage of global land cover changed in one year exactly like other studies (Friedl et al., 2010;Sulla-Menashe et al., 2019).

Change in manuscript:

We have added the corresponding introduction.

L263: “different time periods” the gross change each year or on the difference between the first and the last year in each period?

Response s37:

The annual ratio of the global land cover change area to the global total terrestrial area is plotted in

Fig. 9, but in a form of boxplot organized in a 5-year interval (a) and 10-year interval (b).

Change in manuscript:

We have revised our description.

L267: It's interesting to see a very likely decreasing trend of total LCC area.

Response s38:

Yes, and it was what our results showed.

L270: Fig. 10: The text and subplots in the figure is too small to read. I would suggest to only show the main land cover types and put the others to SI

Response s39:

Thanks for your suggestion.

Change in manuscript:

We have moved the low percentage classes such as shrubland, tundra and snow/ice to the supplementary information part.

L271: why only significant LCC? is it really necessary? why not just sum all?

Response s40:

In our opinion, the statistical test is necessary to lower the uncertainty in the long time-series land cover mapping results, especially for 0.05° such coarse resolution data.

L287: Table 5-10: too detailed, may put into SI and merge these results in a plot with different subplots

Response s41:

Thank you for the advice.

Change in manuscript:

We have reorganized the tables and put some into the supplementary information part.

L299: In In

Response s42:

Thanks for your correction.

Change in manuscript:

We have deleted the extra word.

L304-308: see my comments on greening above

Response s43:

Thanks again.

Change in manuscript:

We have deleted the corresponding part.

L309: need to note the high uncertainty from Fig. 7

Response s44:

Thank you for your valuable comment.

Change in manuscript:

We have added the note in the manuscript.

L310-311: greening again

Response s45:

Thanks again.

Change in manuscript:

We have deleted the corresponding part.

L313: Look at Fig. 10 and 11, significant grassland changed to forest in your dataset in the high latitudes

Response s46:

Yes, it is.

L316-317: why barren land decrease implies the desertification effects?

Response s47:

We may not expressed it clearly. What we mean was the management efforts to the desertification.

Change in manuscript:

We have updated the description.

L321: what is a coupling effect? non-relevant sentence

Response s48:

Thank you for the comment. What we mean was that natural and human factors usually had a significant joint effect on land cover change. Both aspects contribute to making a difference.

Change in manuscript:

We have revised the sentence.

L324 and all below: referring a or b when you report something. Why no explanations on the transitions from grassland to forest, which is the most obvious pattern in your figure

Response s49:

Thanks for your suggestion. This may be related to the shortcomings of the hard classification we adopted. As you pointed out above, the forest may become denser and the land cover class may change. But the interpretations of these phenomena are not the main focus of this paper.

L338: too strong statement. surface greening is not something that you can directly interpreted from LCC.

Response s50:

Thanks for your comment.

Change in manuscript:

We have changed the word.

L345: “natural vegetation” managed forest or pasture are not natural vegetation

Response s51:

Thank you for the comment. We used the wrong word.

Change in manuscript:

We have deleted the word “natural”.

L346: how about reforestation?

Response s52:

Yes, human activities also include reforestation. The focus of this paper is still on data product introduction and evaluation, we will weaken this part of the introduction.

Change in manuscript:

We have revised our description to avoid the ambiguity.

L355: shy subtropical mountain system is also high?

Response s53:

Figure 6 shows the division of eco-regions from FAO, where regions in the orange color belong to subtropical mountain system. Referring to Fig. 5, they overlaps some regions with a relatively high human impact level, such as Spain, central China, east America and South Africa. These regions may bias the overall results.

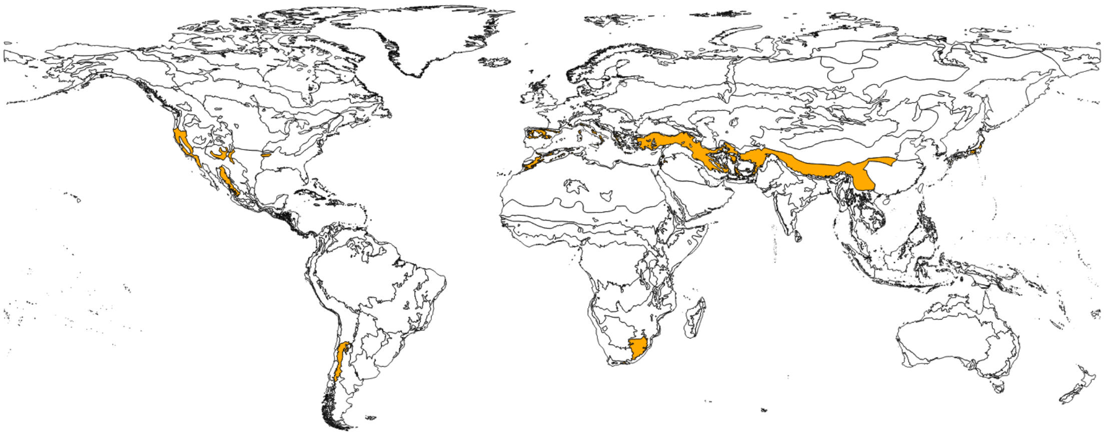

Figure 6 Subtropical mountain system eco-regions from FAO.

L363: Fig. 15 is very misleading with only >0 and <0. Why not give gradient of change?

Response s54:

Thank you very much for your suggestion. Here, we do so because we want to more intuitively reflect the information about where gain or loss occurred.

L365 and below: again, give the subplot title when you describing the results.

Response s55:

Thanks again.

Change in manuscript:

We have added the information as you suggested.

L381: Do you have evidence that global warming will increase vegetation in tropics?!

Response s56:

This part of the analysis is not the focus of this paper.

L383-384: Oil palm plantations are forest or crop in your classification system? I am not sure whether you can distinguish them!

Response s55:

We are sorry for it. In our classification system, oil palm plantations are forest.

Change in manuscript:

We have adjusted the sentence.

L399: Is that partly why you detected forest increase at the expense of grassland?

Response s57:

We are afraid not. Mongolia and Inner Mongolia of China mostly belong to semi-arid regions. The land cover types there should be grassland or barren land. It should have nothing to do with the forest.

L421-423: yes, this is the main defect of this product.

Response s58:

Yes. This is also one of the common problems of hard classification.

L435: This is definitely something that has to be done in this work.

Response s59:

Thank you for your advice. In order to specifically evaluate the magnitude of the errors introduced by our training samples, we randomly selected 500 samples from the training samples for manual interpretation and evaluation, and the assessment accuracy was 92.26%. It shows that the training samples we generate this way are sufficient for our data production.

Change in manuscript:

L441: what about the heterogeneous pixels?

Response s60:

We have added new samples for comprehensive independent accuracy assessment, where heterogeneous samples are also included.

L460: NDVI and LAI increase not equal to forest increase

Response s61:

Indeed it is. NDVI and LAI are features that help forest classification, and we are not strict in saying so here.

Change in manuscript:

We have updated our words.

L455-463: not helping but expose the weakness of the product

Response s62:

Thanks for your comment.

Change in manuscript:

We have deleted the corresponding part.

L464-465: This contradicts that you said forest loss in SE Asia is due to oil palm plantations

Response s63:

We are sorry for it. In our classification system, oil palm plantations are forests.

L466-467: need more explanations to justify the reasons for doing this.

Response s64:

As mentioned above, this is the result of our trade-off. There are too many uncertain factors in remote sensing. In contrast, suppressing some real fluctuations in LCC, and performing post-processing in the time dimension can make data products more reliable, less uncertain and less noisy. And the accuracy improvement brought by change detection illustrates the effectiveness of doing so.

References

Bontemps, S., Defourny, P., Radoux, J., Van Bogaert, E., Lamarche, C., Achard, F., Mayaux, P., Boettcher, M., Brockmann, C., and Kirches, G.: Consistent global land cover maps for climate modelling

communities: Current achievements of the ESA's land cover CCI, Proceedings of the ESA Living Planet Symposium, Edimburgh, 2013, 9-13,
DeFries, R. S., Hansen, M., Townshend, J. R. G., and Sohlberg, R.: Global land cover classifications at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers, International Journal of Remote Sensing, 19, 3141-3168, https://doi.org/10.1080/014311698214235, 1998.
Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., and Huang, X.: MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets, Remote Sensing of Environment, 114, 168-182, https://doi.org/10.1016/j.rse.2009.08.016, 2010.
Gómez, C., White, J. C., and Wulder, M. A.: Optical remotely sensed time series data for land cover classification: A review, ISPRS Journal of Photogrammetry and Remote Sensing, 116, 55-72, https://doi.org/10.1016/j.isprsjprs.2016.03.008, 2016.
Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., and Liu, S.: Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data, International Journal of Remote Sensing, 34, 2607-2654, https://doi.org/10.1080/01431161.2012.748992, 2013.
Li, C., Gong, P., Wang, J., Zhu, Z., Biging, G. S., Yuan, C., Hu, T., Zhang, H., Wang, Q., Li, X., Liu, X., Xu, Y., Guo, J., Liu, C., Hackman, K. O., Zhang, M., Cheng, Y., Yu, L., Yang, J., Huang, H., and Clinton, N.: The first all-season sample set for mapping global land cover with Landsat-8 data, Science Bulletin, 62, 508-515, https://doi.org/10.1016/j.scib.2017.03.011, 2017.
Liu, X., Yu, L., Li, W., Peng, D., Zhong, L., Li, L., Xin, Q., Lu, H., Yu, C., and Gong, P.: Comparison of country-level cropland areas between ESA-CCI land cover maps and FAOSTAT data, International Journal of Remote Sensing, 1-15, https://doi.org/10.1080/01431161.2018.1465613, 2018.
Plaza, A., Martínez, P., Pérez, R., and Plaza, J.: Spatial/spectral endmember extraction by multidimensional morphological operations, IEEE transactions on geoscience and remote sensing, 40, 2025-2041, https://doi.org/10.1109/TGRS.2002.802494, 2002.
Song, X.-P., Hansen, M. C., Stehman, S. V., Potapov, P. V., Tyukavina, A., Vermote, E. F., and Townshend, J. R.: Global land change from 1982 to 2016, Nature, 560, 639, https://doi.org/10.1038/s41586-018-0411-9, 2018.
Sulla-Menashe, D., Gray, J. M., Abercrombie, S. P., and Friedl, M. A.: Hierarchical mapping of annual global land cover 2001 to present: The MODIS Collection 6 Land Cover product, Remote sensing of environment, 222, 183-194, https://doi.org/10.1016/j.rse.2018.12.013, 2019.
Wang, C., Gao, Q., Wang, X., and Yu, M.: Spatially differentiated trends in urbanization, agricultural land abandonment and reclamation, and woodland recovery in Northern China, Scientific reports, 6, 37658, https://doi.org/10.1038/srep37658, 2016.
Zhang, B., Sun, X., Gao, L., and Yang, L.: Endmember extraction of hyperspectral remote sensing images based on the ant colony optimization (ACO) algorithm, IEEE transactions on geoscience and remote sensing, 49, 2635-2646, https://doi.org/10.1109/TGRS.2011.2108305, 2011.