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

General comment:

By fusing multiple existing geo-spatial datasets, the main work of this manuscript is to generate an annual dynamic product (spatial resolution: 0.5°) addressing seven kinds of land-covers (i.e., cropland, forest, grassland, shrub-land, tundra, barren land and snow/ice) from 1982 to 2015. With eye on the current existing datasets (i.e., from the perspective of classification system, period of time, and spatial/temporal resolution) the contribution is quite limited. In view of the rationality of technique and accuracy assessment, current version calls for serious revision before publication. In view of the analysis conducted on the dynamic map, rare novel findings can be captured.

Response 1:

Thanks for the comment. First of all, this is a paper describing a unique data product. It is not 0.5 degrees in resolution but 5 km. Since it is about land cover data product, it is not our attention to make novel discoveries. The purpose here is mainly to present a data set that does not exist anywhere before, for its annual frequency, 34 years long duration and high accuracy. The classification system does cover more than 90% of the land area. We did not include water, wetland, and impervious areas because wetland is extremely dynamic (more frequent than the yearly scale), water excluded from the input data source, and impervious areas already processed using more accurate source of data (e.g., annual Global Artificial Impervious Area maps, (Gong et al., 2020)). The accuracy assessment has been further improved using additional collection of test samples. We also compared our results with other data products and found that our results are superior.

Specific comments:

There are several global datasets with more rigorous production process have existed. 1) The 1992-2018 annual 300m global land-cover data (https://www.esa-landcover-cci.org/?q=node/197) with more detailed classification scheme have been released. Since the proposed product has no accuracy assessment on the annual maps from 1982-1991, it cannot be argued that the proposed work have longer period of time.

Response 2:

Thanks for your comment. We agree that ESA-CCI products have higher spatial resolution and more detailed classes. However, products with different resolution have different application purposes. In many studies, it is only necessary to use coarse-resolution land cover data, such as our 0.05 ° data, which can be used in Earth system modeling.

For Earth system modeling purposes, the 10 land cover classes mentioned in our response at the beginning are sufficient. Among the ten classes, except for wetland, impervious area, and water that occupy less than 10% of the entire land area on Earth. In the meantime, water and impervious areas can be individually obtained. Wetland is highly dynamic requiring additional types of remotely sensed data. Considering the separability and identifiability of the land cover classes under the 5 km spatial resolution, we adopted a classification system of 7 classes.

In this revision, we collected new independent test samples and performed accuracy assessment for the period of 1982-1991. In addition, we have compared our products with ESA-CCI and MODIS-based land cover data products and FAOSTAT data. The results show that our products have good reliability.

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, in order 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.

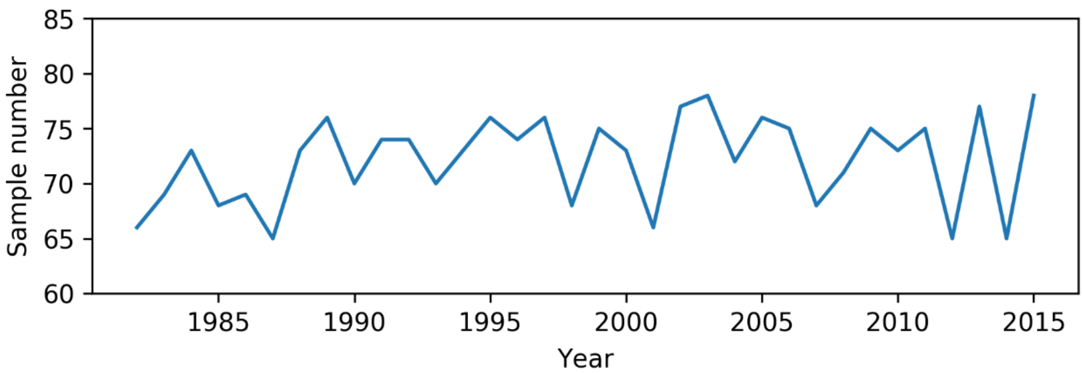

Figure 1: The temporal distribution of the newly collected test sample.

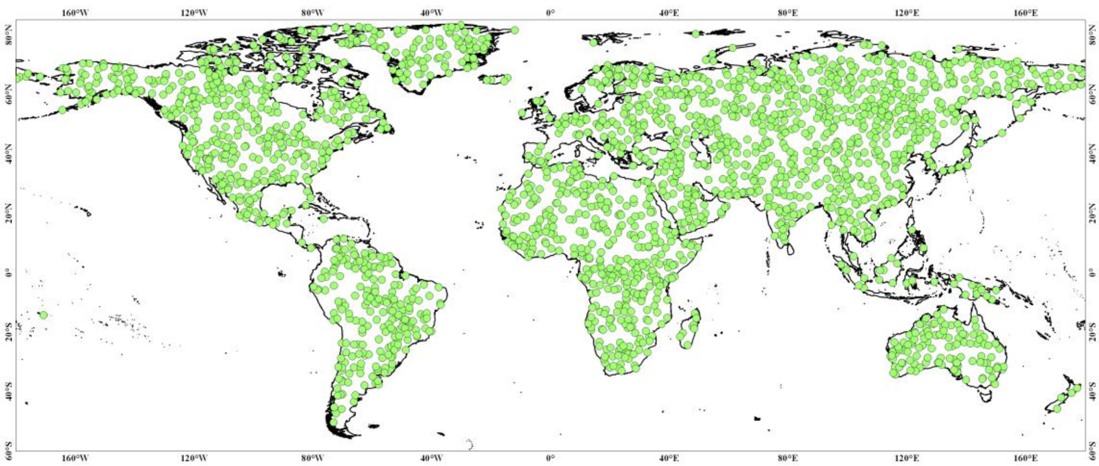

Figure 2: The geographical distribution of random test sample.

The new assessment result is shown in Table 1 and Table 2. 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 1: 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 2: 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%

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 3: 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 4: 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 5: 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 shows an inter-comparison with MODIS-based products, Figure with ESA-CCI products and Figure 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 during 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 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.

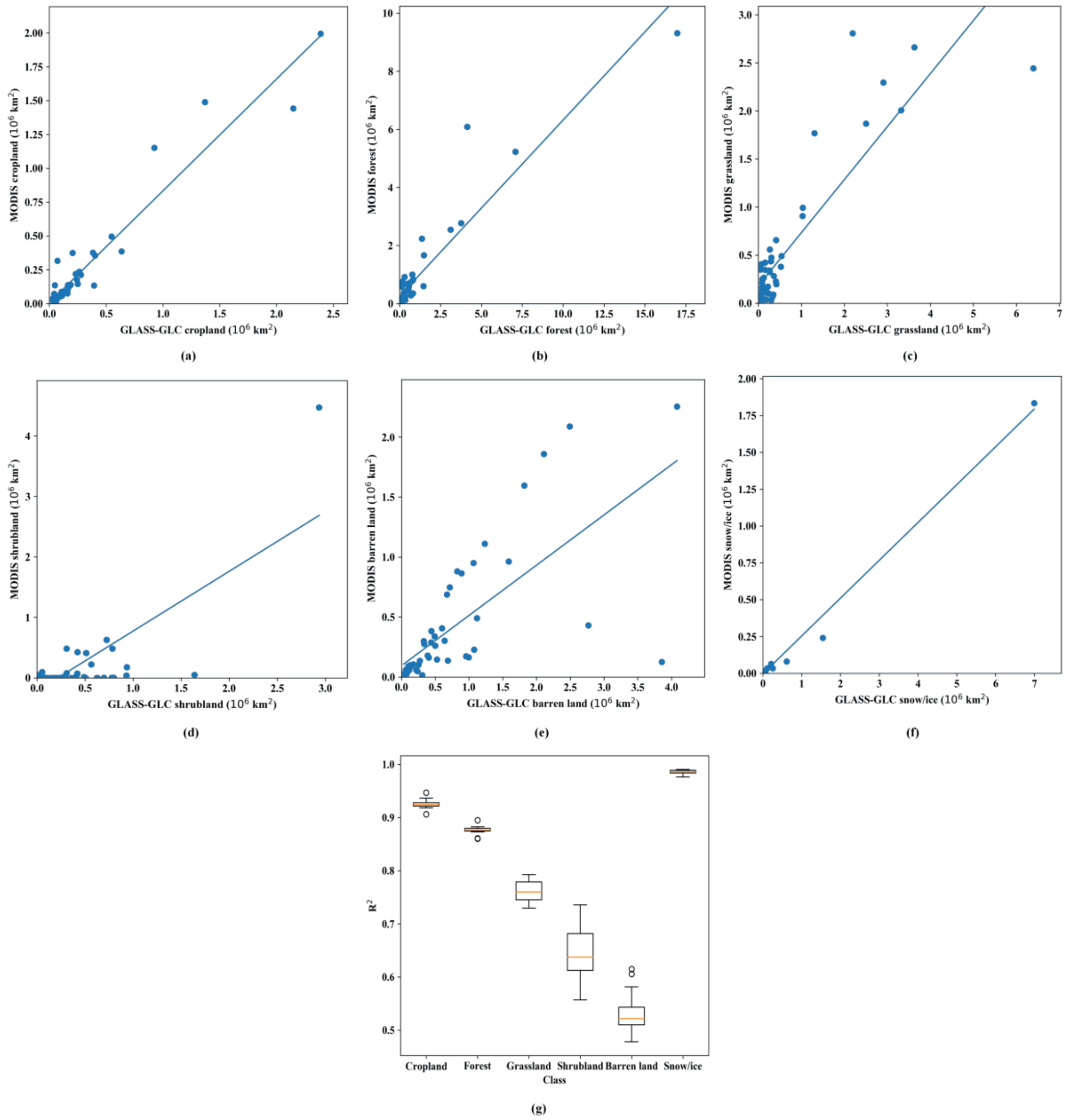

Figure 3: Inter-comparison with the MODIS-based land cover product, (a) cropland circa 2015, (b) forest circa 2015, (c) grassland circa 2015, (d) shrubland circa 2015, (e) barren land circa 2015 and (f) snow/ice circa 2015; (g) mean R2 of the annual linear fit lines for all years (2001-2015).

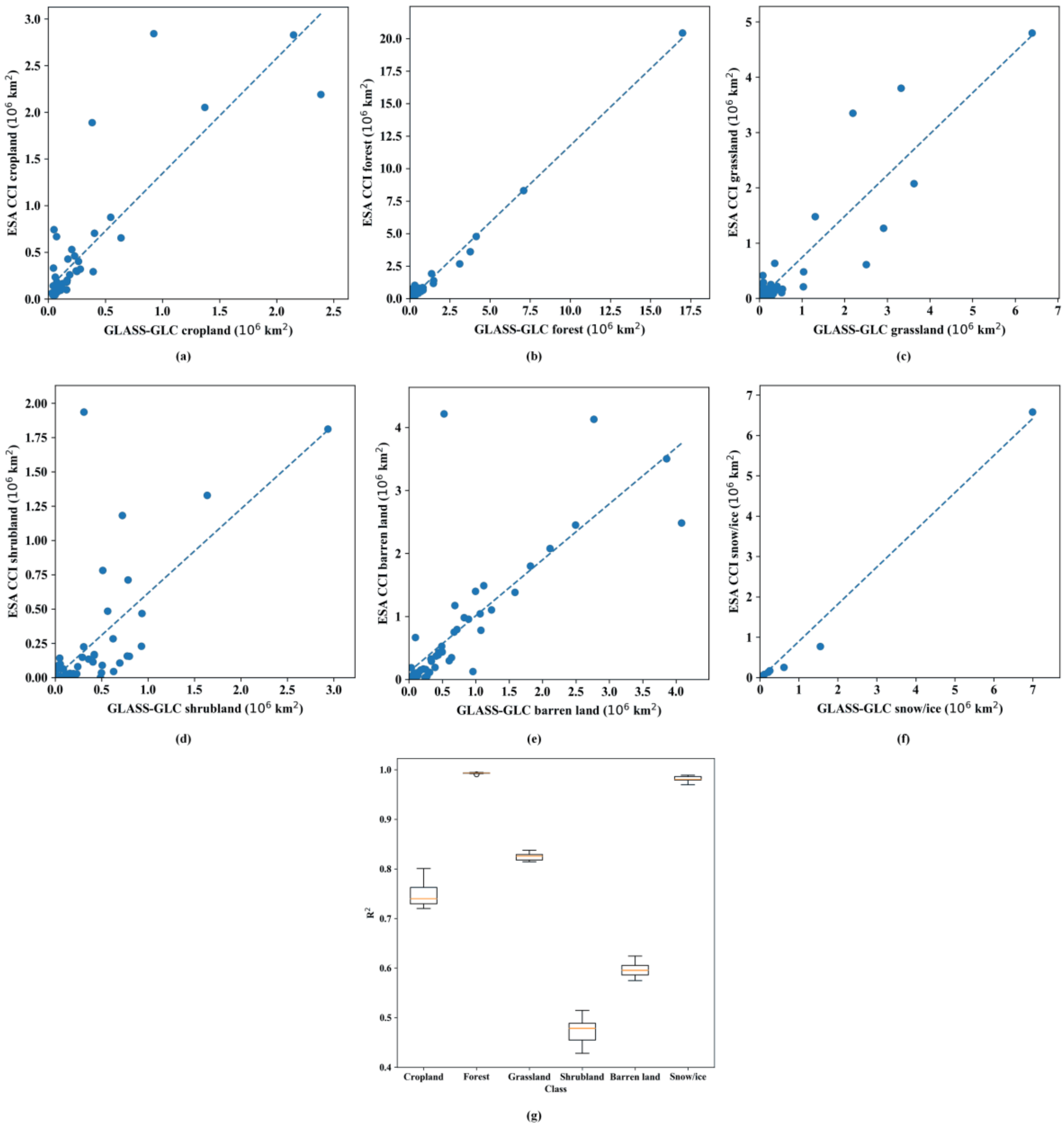

Figure 4: Inter-comparison with the ESA-CCI land cover product, (a) cropland circa 2015, (b) forest circa 2015, (c) grassland circa 2015, (d) shrubland circa 2015, (e) barren land circa 2015 and (f) snow/ice circa 2015; (g) mean R2 of the annual linear fit lines for all years (1992-2015).

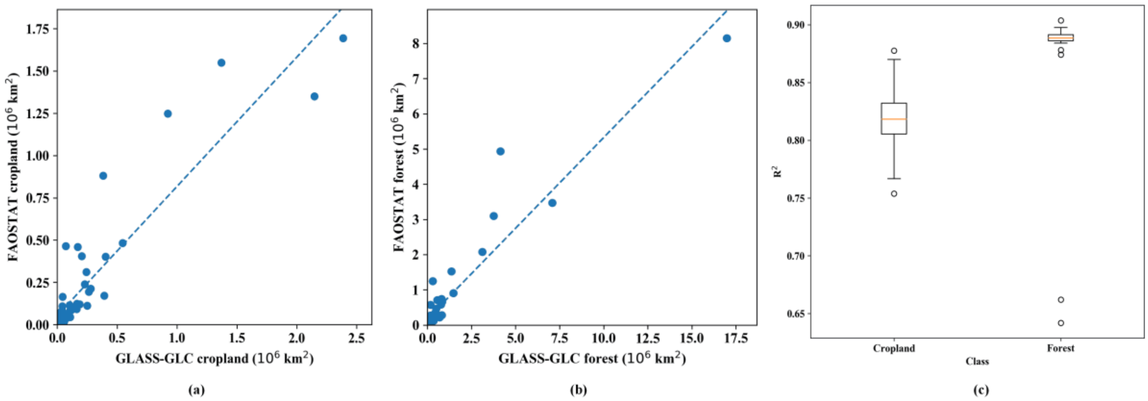

Figure 5: Inter-comparison with the FAOSTAT data set, (a) cropland circa 2015, (b) forest circa 2015, (c) mean R2 of the annual linear fit lines of cropland for years 1982-2015 and forest for years 1990-2015.

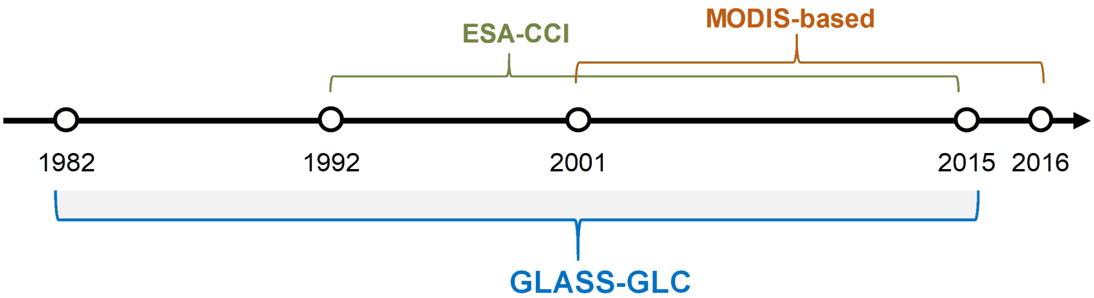

Figure 6: The duration of different land cover products, including GLASS-GLC, MODIS-based land cover products and ESA-CCI land cover products.

Change in manuscript:

We have added the new accuracy assessment results and data inter-comparison results to help show the reliability and effectiveness of GLASS-GLC.

2) The annual VCF products from 1982- 2016 have the same spatial resolution, and very similar classification scheme with the proposed work (from 1982-2015). Although the VCF product is missing in 1994 and 2000, the proposed work just directly use the data-source around the adjacent year, which cannot be viewed as a noticeable contribution. Meanwhile, since the proposed work also introduce VCF in the supervised classification, the analysis on the dynamic map is somewhat similar to this existed study (Song et al 2018a) , but more superficial.

Response 3:

Thank you for your comment. VCF is a quantitative variable. VCF data products mainly reflect vegetation cover information. Our land cover classes include multiple types of nominal variables. Again VCF and land cover information have different purposes of applications.

Here, we introduce VCF as a priori information to assist in land cover classification. VCF data is missing for two years, but this will not greatly affect the classification results. The auxiliary or supplementary data for classification and interpretation do not need to be perfect. They do not need to be in the same period or at the same resolution. As long as there is supplementary information, it will work, such as in the four-dimensional variational data assimilation.

For the analysis part of the land cover classification, the results are similar to those obtained from the analysis of VCF, which also confirm the objectivity and correctness of VCF analysis. But it is worth pointing out that our products can analyze many more detailed classes, so we can also draw some different conclusions.

Considering that the type of this paper is a data paper, our main focus is on the description of the production methods and quality control of data products, and the comparison and analysis of data quality and accuracy. More in-depth LCC analysis is out of the scope of this study.

Technical corrections

1. It is ridiculous to produce training and test set from a same product and in a same manner. In addition, it is unacceptable to conclude the applicability of the long-time period product by assessing the accuracy on only the 2015 land-cover mapping result.

Response 4:

Thank you for your comment. There may be some flaws in the way we evaluate accuracy.

Taking this into consideration, in addition to the accuracy assessment of samples taken from the FROM-GLC_v2 product, samples from FLUXNET site data are also given for independent accuracy assessment. The assessment results are shown in Table 3. The overall accuracy of GLASS-GLC reached 82.10% in 2015.

In addition, as described in response 2 above, we conducted a new independent sample test (OA=82.81%) and a comparison of multiple products (land cover products from MODIS and ESA-CCI, and FAOSTAT data), which also proved the reliability of our products.

Change in manuscript:

We have added the new accuracy assessment results and data inter-comparison results to help show the reliability and effectiveness of GLASS-GLC.

2. How to project the 30m FROM-GLC_v2 to mapping scale? How to deal with the mixed sample?

Response 5:

Thank you for your question. As the paper says, we projected the results of FROM-GLC_v2 according to the principle of majority. That is, the land cover class that accounts for the largest proportion in each grid is used as the land cover class label under the 0.05 ° grid. Generating coarse-resolution samples from high-resolution products as such is actually a common practice (Wang et al., 2016;DeFries et al., 1998).

For mixed samples, we also use the majority principle to give labels. Although percentage information is more suitable for dealing with mixed pixels, our goal here is hard classification, and we cannot avoid only doing so. This is also a problem that arises in hard land classification studies.

Considering the cost constraints, we have adopted this method of generating new samples even though it will bring some errors when producing coarse resolution samples from FROM-GLC_v2. However, the "stable classification with limited sample" theory (Gong et al., 2019) supports our approach to some extent. The theory shows that under its experimental conditions, even if 20% of the wrong samples are introduced, the classification accuracy is reduced by 1%, and it can still be stable (Fig. 7).

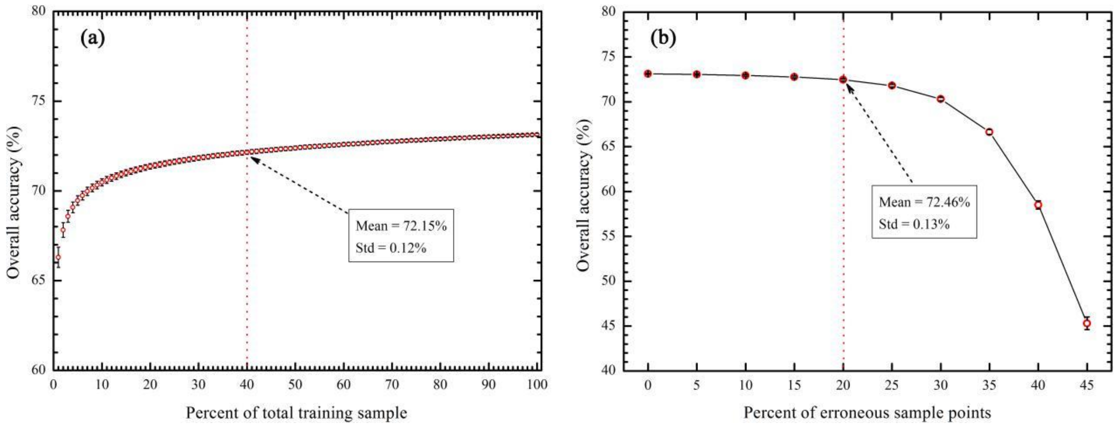

Figure 7: Sample robustness to size reduction and errors in sample. a. As sample size increases, the accuracy quickly reaches a plateau. b. As the impurity percentage of sample increases the accuracy decreases. In both cases, the 1000 times random drawing of sample points produced very stable overall classification accuracies with most standard deviations much lower than 0.5%. (Gong et al., 2019)

The newly added results of accuracy assessment have also confirmed that the samples produced in this way can meet the production needs.

3. There is no sample accuracy assessment on the produced training sample set. Please note that the accuracy of the FROM-GLC_v2 is not high enough to work as training sample.

Response 6:

Thank you for your point. The classification accuracy of FROM-GLC_v2 will surely have some impact on our results. However, FROM-GLC_v2 has been published, and it has a detailed accuracy assessment, with an OA of 73.13%. There is some complicated relationship for the use of higher resolution land cover data in producing lower-resolution land-cover products. Since there is a scaling down which requires aggregation of high-resolution land cover results. This often acts as an averaging effect that improves the accuracy in the area to some extent. Even if there is no accuracy increase during the scaling down process, the 73% accuracy would not cause a large accuracy decrease as can be seen from the figure in the right-hand side of Fig. 7.

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:

We have added the accuracy assessment results on training samples.

4. When mapping the land-covers decades year ago, the suitability of the samples collected (mainly from 2013-2015) should be evaluated.

Response 7:

Thanks for your advice. We agree that, in the early years, the percent of land cover change may be relatively large. However, global land cover will not change by more than a few percents for decades. And these changes are primarily in urban and urban-rural fringe areas. The outdatedness of samples will not affect much of our accuracy assessment.

Concerning the reliability of sample migration, the “stable classification with limited sample” theory is specifically discussed (Gong et al., 2019).

In this study, the concept of a stable classification is defined. They use this concept to approximately determine how much reduction in training sample and how much land cover change or image interpretation error can be acceptable. If the mean accuracy of multiple runs of a classifier trained with a random drawing of a certain percentage of sample points from the total sample is within 1% of what can be achieved with the total sample set, we regard the obtained classification result “stable”. The 1% threshold is empirically chosen based on the fact that a loss of overall accuracy in 1% shall not significantly impact the application of a global land cover map.

Tens of millions of experiments suggest that it is possible to use 60% fewer sample points and even the land cover changed by 20% or the training sample contains 20% errors, we are still able to achieve “stable” classification with the random forest classifier in global land cover mapping. This conclusion well supports the effectiveness of our sample transfer method. Even for decades, it is difficult for global land cover to change by more than 20%. Therefore, the proportion of error samples we introduced in the early years will not exceed 20%, and the classification results are still reliable and effective.

Another recent study (Huang et al., 2020) also devoted to migrating training samples to early years. They developed an automatic training sample migration method, which can successfully migrate training samples in 2015 to 2000. These studies prove the effectiveness of sample migration and provide potential solutions to resolve the problem of lack of training samples for dynamic global land cover mapping efforts.

Besides, to verify the temporal accuracy of our products, as mentioned above, we have independently collected test samples from different years and tested the accuracy of our products, with an accuracy of 82.81%. What's more, the inter-comparison results with other data have also confirmed the validity of our data using the 2015 sample for many years.

Reference

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.

Gong, P., Liu, H., Zhang, M., Li, C., Wang, J., Huang, H., Clinton, N., Ji, L., Li, W., Bai, Y., Chen, B., Xu, B., Zhu, Z., Yuan, C., Suen, H. P., Guo, J., Xu, N., Li, W., Zhao, Y., Yang, J., Yu, C., Wang, X., Fu, H., Yu, L., Dronova, I., Hui, F., Cheng, X., Shi, X., Xiao, F., Liu, Q., and Song, L.: Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017, Science Bulletin, 64, 370, https://doi.org/10.1016/j.scib.2019.03.002, 2019.

Gong, P., Li, X., Wang, J., Bai, Y., Chen, B., Hu, T., Liu, X., Xu, B., Yang, J., and Zhang, W.: Annual maps of global artificial impervious area (GAIA) between 1985 and 2018, Remote Sensing of Environment, 236, 111510, https://doi.org/10.1016/j.rse.2019.111510, 2020.

Huang, H., Wang, J., Liu, C., Liang, L., Li, C., and Gong, P.: The migration of training samples towards dynamic global land cover mapping, ISPRS Journal of Photogrammetry and Remote Sensing, 161, 27-36, https://doi.org/10.1016/j.isprsjprs.2020.01.010, 2020.

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.

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.