Referee #1

General comments:

The issues have been addressed in the manuscript. My suggestion is Accept

Response:

We appreciate your considerable comments and decision. We have carefully revised the manuscript and made technical corrections to further improve the manuscript's quality.

Referee #3

We thank the reviewer and editor for a thoughtful and thorough review of our manuscript (ESSD-2023-87) and for giving the decision on minor revisions. We have carefully addressed all the comments and revised the manuscript. The suggestions and comments are listed in **bold** type. The modified words or materials are marked as blue color in the revised manuscript. The item-by-item responses to all comments are listed below.

General comments:

The manuscript takes on the very substantial challenge of developing a national land-scale land cover of China with 1-meter spatial resolution. The newly developed very-high-resolution (VHR) land cover data was established using a deep learning-based framework and open-access data including global land-cover (GLC) products, open street map (OSM), and Google Earth imagery. In general, the revised manuscript is well structured, and authors did a lot of work in addressing the two reviewer's comments, especially the data validation. However, the texts about validation and comparison still need to be improved. Some detailed suggestions and comments were provided below. Overall, this paper could be publishable after revision.

Response:

We are grateful for your careful reading, and for giving us constructive comments. According to the suggestions and comments, we have carefully considered all of them and tried our best to improve the manuscript before it could be published especially the validation and comparison parts.

Suggestions and comments:

(1) Line 187 and Table 1. In this manuscript, the authors defined a classification system, including 11 land cover types. Though you tried to match the classification definition between FROM_GLC10, ESA_GLC10, ESRI_GLC10 and SinoLC-1. However, the land cover class definition between the 10-m resolution is also different. For example, the tree cover in ESA_GLC10 includes any geographic area dominated by trees with a cover of 10% or more. Forest in FROM_GLC10 limits tree cover percentage classification to >15%, limits tree height classification to >3 m (Gong et al., 2012, IJRS). Trees in ESRI_GLC10 is defined as any significant clustering of tall (~15 feet or higher) dense vegetation, typically with a closed or dense canopy[1]. Other land use types should also have large differences. So, how do you consider such differences when selecting training and validation samples? Besides, a clear definition for each land use type in SinoLC-1 should be given in a table.

Response:

We appreciate the questions and suggestions from the reviewer. We agree that a clear definition for each land-cover type in the SinoLC-1 product is important for the users and downstream applications. The item-by-item responses to this comment are listed below.

During the selection of training and validation samples, we considered three main issues and tried to resolve them by conducting a weakly supervised strategy in the mapping framework. Firstly, as the reviewer mentioned, there is a difference between the type definition of the utilized global 10-meter products. Secondly, there are resolution gaps between the imagery and the 10-meter products. Thirdly, there is a temporal mismatch between the three utilized 10-meter products. The above-mentioned problems can be all regarded as "noisy label issues", where the training labels have unreliable samples. Figure R1 shows a case of the label generation process that takes the intersection parts of three products. Although there are label noises caused by the mismatched type definition, spatial resolution, and temporal resolution, the reminded samples shown in Figure R1 (e) are still relatively reliable. Based on such intersection labels, the result of the SinoLC-1 shown in Figure R1 (f) is accurate and consists of the 1-meter imagery.

[1] https://www.arcgis.com/home/item.html?id=d3da5dd386d140cf93fc9ecbf8da5e31

For giving a clearer explanation to the readers, firstly, we added a description of type unification in line 20 to descript the differences in type definitions among the three 10-meter products. Secondly, we supplemented the clear definition of each land-cover type of the proposed SinoLC-1 in Table R1 (Table 2 of the revised manuscript). By combining the actual situation of the proposed SinoLC-1 product and the information from the China Ministry of Natural Resources, the definitions including detailed descriptions and examples were given to better explain each land-cover type.

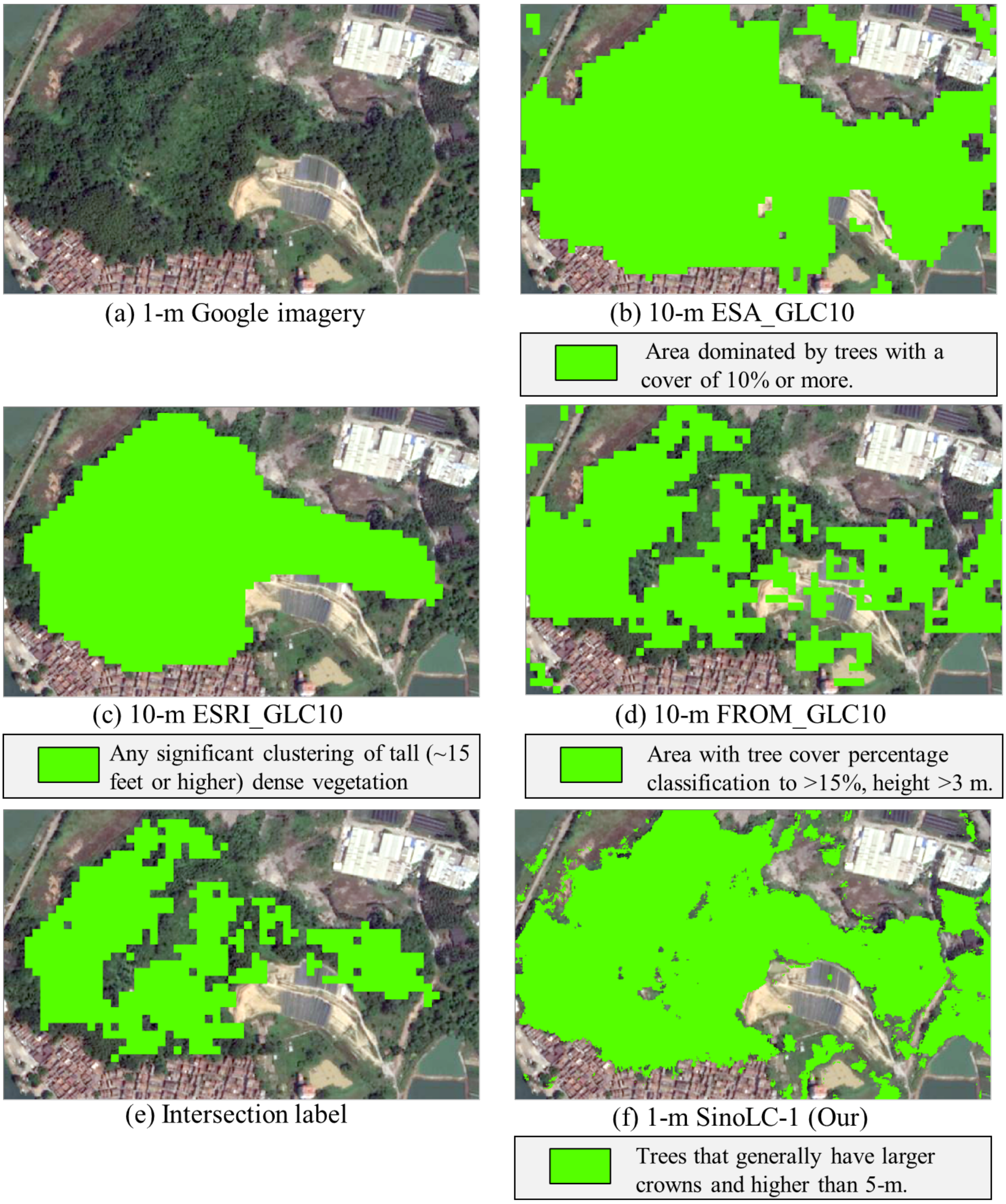

Figure R1. A case (Tree cover) of the label generation process of the SinoLC-1.

Table R1. The definition, value, and color of each land-cover type of the SinoLC-1

Land-cover type	Definition	Value	Color	
Tree cover	Areas covered by trees generally have larger crowns and are higher than 5 meters. It can be sparse arbors or clustered forests which include evergreen forests, mixed forests, artificial forests, bamboo groves, etc.	2	(0, 100, 0)	
Shrubland	Areas covered by clusters of shrubs with a height below 5 meters.	3	(255, 190, 35)	
Grassland	Areas covered by low herbaceous plants. It generally includes natural grasslands with a fractional vegetation coverage greater than 5, rangeland with tree canopy density less than 0.3 or shrub canopy density less than 0.4, urban's vacant land dominated by grass, and other artificial grasslands.	4	(233, 255, 190)	
Cropland	The arable land and human planted crops not at tree height including upland crops (e.g., wheat, corn, potatoes, and cotton) and irrigated crops (e.g., paddy filed, lotus root, and water spinach).	5	(255, 235, 175)	
Building	Human-made structures and homogenous impervious surfaces including industrial, residential, commercial areas, and construction sites. It is generally located in urban and rural areas with high human activities.	6	(255, 170, 0)	
Traffic route	Areas constructed according to certain technical standards and equipped with necessary transportation facilities, including railways, highways, urban/rural roads, and pipelines.	1	(255, 0, 0)	
Barren and sparse vegetation	Areas covered by sparse vegetation or bare land covered by sand, gravel, or rocks, including mountains without dense vegetation and snow cover, deserts, grasslands degraded by drought, and wasteland in urban/rural areas with sparse or no vegetation.	7	(180, 180, 180)	
Snow and ice	Areas covered by large-scale permanent snow or ice, including glaciers and permanent snowpack in mountain areas or high latitudes.	8	(240, 240, 240)	
Water	Areas covered by water for a long period, including oceans, naturally formed water bodies (e.g., lakes, rivers, and runoff), and artificially formed water bodies (e.g., reservoirs, canals, water conservancy facilities, ponds, and aquaculture farms).	9	(0, 100, 200)	
Wetland	Areas with perennial or seasonal water accumulation and vegetation growth. It includes forest/shrub/grass swamps, peatlands, mudflats, mangroves, and coastal/inland tidal flats.	10	(0, 150, 160)	
Moss and lichen	Surfaces or rocks attached by moss or tiny lichen plants.	12	(250, 230, 160)	

(2) The Google Earth (GE) images were used for land cover classification, but it is like a figure with R, G, and B bands rather than the satellite image with multiple optical bands. Though a deep machine learning method was used to extract the features, the classification process was conducted at the pixel level. Object-based methods may have more advantages than pixel-based methods for VHR image classification. Why didn't choose the object-based classification method? As it should be, the pixel-based method can also generate accurate land cover. But a post-processing step should be included because the pixel-based method would always generate many fragmented patches.

Response:

Thank you for your thoughtful consideration and constructive comments. We would like to respond to them in three aspects:

Firstly, by considering the very large-scale land-cover mapping process (i.e., the entire China covering 9,600,000 km^2), we designed a deep learning framework with a weakly supervised strategy to deal with the resolution mismatched issue between the training labels and imagery. On one hand, deep learning technique has been adopted in large-scale land-cover mapping in many cases where most of the deep learning frameworks are end-to-end trainable with pixel-based classification head (Li et al., 2022; Liu et al., 2023). However, most of the Object-based Image analysis (OBIA) methods are integrated into geographic software (e.g., ArcGIS and eCognition) which require human intervention and hand-craft labeled data[2]. On the other hand, the proposed framework in our manuscript has high efficiency in training and inference which shortens the time cycle for nationwide land-over mapping. However, as shown in Figure R2, the OBIA method is generally time-consuming in object generation and the parameters for generating objects often need to be carefully selected according to the situation. In conclusion, by considering the attributes of training data, the long span of the mapping areas, various landforms throughout China, and the efficiency of land-cover mapping, we chose the deep learning framework with a regular pixel-based classification head in the production of SinoLC-1.

[2] https://gisgeography.com/obia-object-based-image-analysis-geobia/

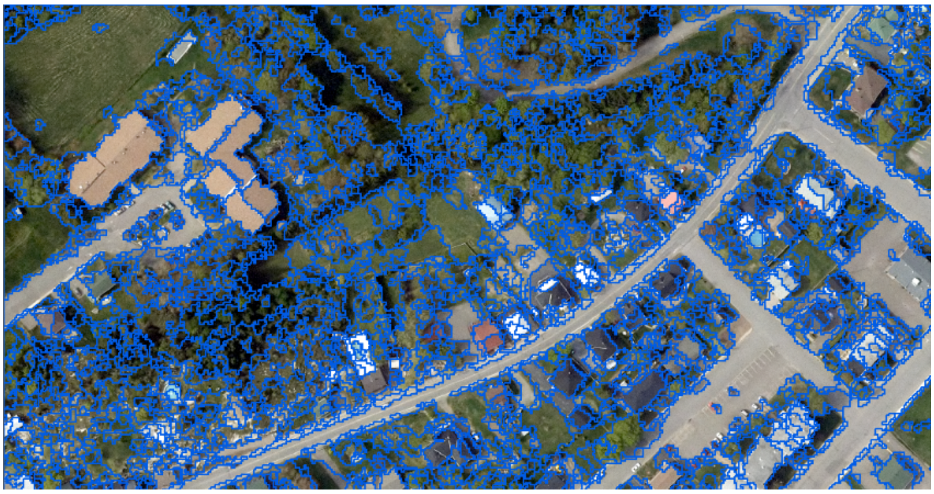

Figure R2. ArcGIS segmentation using the Segment Mean Shift algorithm.

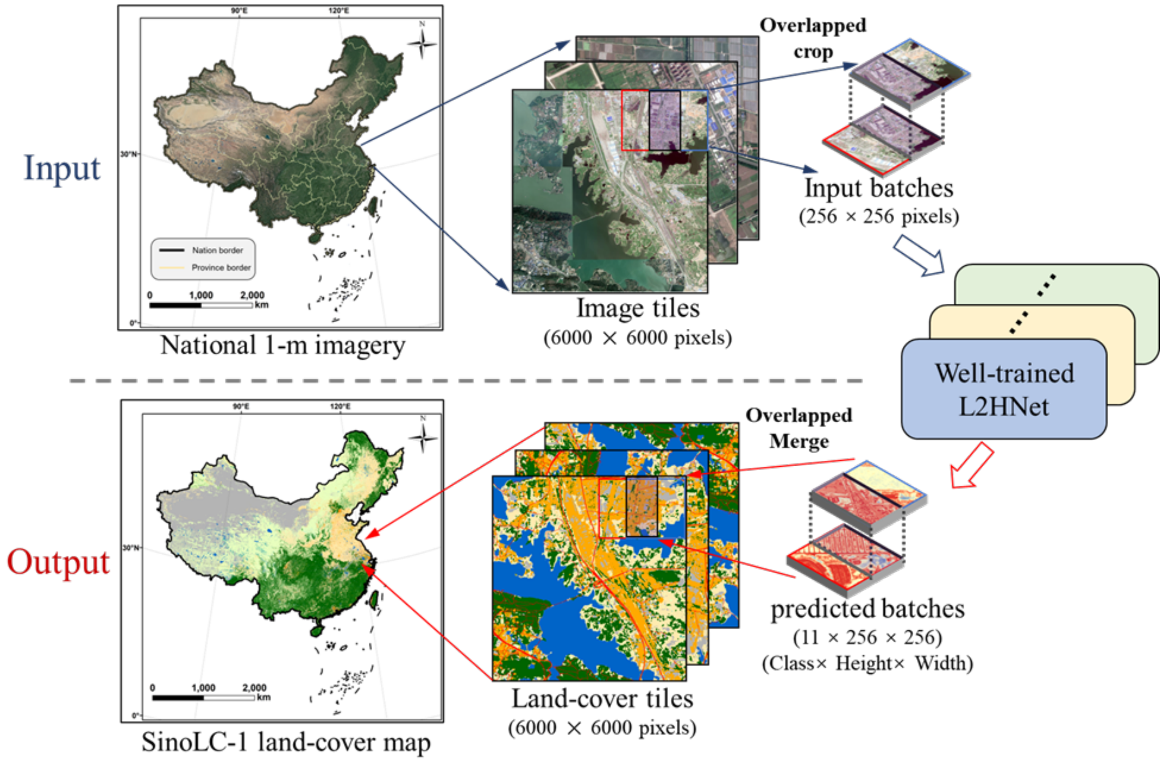

Figure R3 Demonstration of the mapping and merging for producing SinoLC-1.

Secondly, as the mapping and merging process shown in Figure R3, the large-scale images need to be cropped into many tiles (6000×6000 pixels) and batches (256×256 pixels) during the training and inference process. The edges adjacent to image batches usually require taking the predicted average or majority of two adjacent batches to improve the results on the edge. However, it is difficult to the object-based classification since there are differences in the object boundaries of adjacent batches, which bring batches' edges mismatched during the large-scale land-cover mapping process.

Thirdly, we agree that the post-processing step can remove the fragmented patches inferred by the pixel-based classification methods. According to your suggestion, we have tried to conduct the post-processing in many regions by fusing multimodel results and enhancing typical land-cover types (Li et al., 2021). As shown in Figure R4, we used two-step post-processing to improve the final results, and most of the salt and pepper look in the classification result was removed. We will attempt to conduct the post-processing throughout the product and continuously update the published version in the future.

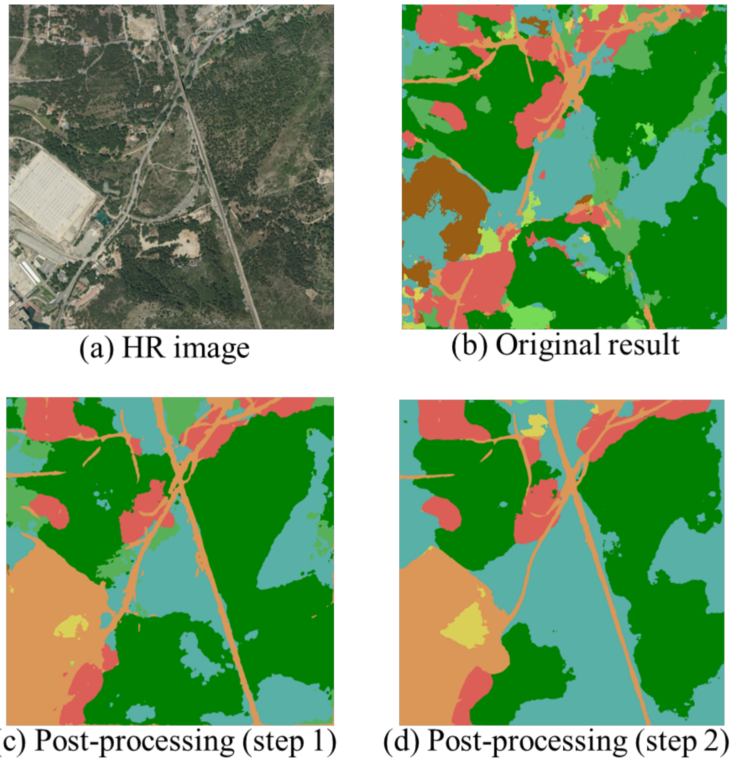

Figure R4. An example of different steps of the post-processing results

The cited references of this response are as follows:

Li, Z., Zhang, H., Lu, F., Xue, R., Yang, G., & Zhang, L. (2022). Breaking the resolution barrier: A low-to-high network for large-scale high-resolution land-cover mapping using low-resolution labels. ISPRS Journal of Photogrammetry and Remote Sensing, 192, 244-267. https://doi.org/10.1016/j.isprsjprs.2022.08.008

Liu, S., Wang, H., Hu, Y., Zhang, M., Zhu, Y., Wang, Z., ... & Wang, F. (2023). Land Use and Land Cover Mapping in China Using Multi-modal Fine-grained Dual Network. IEEE Transactions on Geoscience and Remote Sensing. 10.1109/TGRS.2023.3285912

Li, Z., Lu, F., Zhang, H., Yang, G., & Zhang, L. (2021, July). Change cross-detection based on label improvements and multi-model fusion for multi-temporal remote sensing images. In 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS (pp. 2054-2057). IEEE. 10.1109/IGARSS47720.2021.9553120

(3) Line 350-351. "The tree canopies and dense vegetation are mainly in the southern part and the northeast border of China". Here should be tree cover (defined in Table 1) rather than tree canopies and dense vegetation. Many different words were used to describe the tree cover in the texts and figures. For example, tree cover is labeled as tree in Figure 4b, Figure 6, Trees in Figure 8, Forest in Figure 13, Tree canopy in Figure 14, respectively. It will confuse readers and should be consistent in the text and figure legend.

Response:

Thanks for the comments and corrections. We have checked the whole manuscript and corrected all descriptions of land-cover types which include the legends of Figure 4b, Figure 6, Figure 8, Figure 13, Figure 14, Figure 15, Figure 21, and Figure 22.

(4) Line 488. Northern China, including Beijing, Tianjin, Hebei, Shanxi, and Inner Mongolia, have the lowest O.A. (lower than 70%) because the longitude span of the region is very wide, and the landscapes are diverse and various. For Inner Mongolia, the explanation looks good. But it is unreasonable for Beijing, Tianjin, Hebei, and Shanxi. I think the reason can be found in the confusion matrix of these provinces, especially which land cover type has the lowest accuracy.

Response:

Thank you for the constructive comments. To better explain and analyze the classification results in Northern China, we carefully checked the confusion matrix of Beijing, Tianjin, Hebei, and Shanxi and summarized the confused land-cover types which cause the low accuracy in these regions. For Beijing, most of the misclassified samples are (1) the confusion between "Tree cover" and "Grassland"; (2) the confusion between "Building" and "Traffic route". For Tianjin, most of the misclassified sample is the confusion among "Cropland", "Building", and "Traffic route". For Hebei, most of the misclassified samples are (1) the confusion between "Tree cover" and "Grassland"; (2) the confusion between "Cropland" and "Grassland". For Shanxi, most of the misclassified samples are (1) the confusion among "Tree cover", "Grassland", and "Cropland"; (2) the confusion between "Building" and "Traffic route"; (3) the confusion between "Cropland" and "Barren & sparse vegetation". Detailed explanations were added in Section 4.3 "Quantitative analysis and accuracy assessment" to analyze the reason causing low accuracy in Northern China.

(5) Section 4.2.2, Quantitative comparison with other land-cover products. You did the same validation work as that in section 4.3.1 rather than data comparisons between SinoLC-1 and the other five land cover datasets. I suggest moving this section to section 4.3.1 or doing a data comparison like section 4.3.2.

Response:

Thanks for your constructive suggestions. We agree that changing the section order of quantitative comparison can better improve the structure and readability of the manuscript. We have moved Section 4.2.2 “Quantitative comparison with other land-cover products” to Section 4.3 “Quantitative analysis and accuracy assessment” which can better support the analysis of Section 4.3.1 “Pixel-level sample validation”.

(6) Figure 22 shows that there are significant differences in some provinces between SinoLC-1 and NLRS data. For example, the differences in tree cover area of Chongqing and Hainan are larger than 40%. Some explanations for such large differences should be given. Moreover, it is also looks weird to combine the grassland and barren into one type for comparison.

Response:

Thank you for your constructive comment. For the first suggestion, we carefully checked the qualitative results and the statistical results of the SinoLC-1 and NLRS data. The corresponding analysis, which aims at explaining the large differences and misestimation of the statistical comparison, was supplemented in Section 4.3.3 “Statistical-level validation”. Especially, according to the concerns in the comments regarding the results of Hainan and Chongqing, we demonstrated the qualitative comparison and the statistical comparison of the SinoLC-1 shown in Figure R5 and Figure R6. From the statistical comparison results, the SinoLC-1 of Hainan and Chongqing has a high overestimation of “Tree cover” and an underestimation of “Cropland”. From the qualitative comparison results, the SinoLC-1 of Hainan and Chongqing Provinces are more consistent with the 10-m ESA_GLC10 where the “Tree cover” occupies more large areas than the FROM_GLC10 and ESRI_GLC10. In general, Hainan and Chongqing Provinces have a high proportion of “Tree cover” in practice, and the labels generated for model training retain massive samples of “Tree cover” in these two areas, which led to the model overfitting and overestimating the types of “Tree cover” (underestimating the type of “Cropland”).

For the second question, we combined the land-cover types of “Grassland” and “Barren and sparse vegetation” by considering that the type of “Barren and sparse vegetation” also includes the areas with low and sparse grass cover. Besides, we carefully checked the document ‘Detailed Rules for the Recognition of Land Classification in the Third National Land Survey (2019)’ at the website of the Ministry of Natural Resources of the People's Republic of China (https://m.mnr.gov.cn/zt/td/dscqggtdc/zl/201906/P020190604539900543194.pdf). In the classification system of NRLS data, there is no land-cover type used to describe bare land with sparse vegetation. However, as shown in Table R2, the land-cover type of “Grassland” includes many subcategories that descript sparse vegetation coverage areas. Furthermore, in the classification system of SinoLC-1, the “Barren and sparse vegetation” also includes the grasslands degraded by drought and wasteland in urban/rural areas with sparse or no vegetation. In general, based on the classification system of NLRS data and SinoLC-1, we combined the “Grassland” and “Barren and sparse vegetation” of the SinoLC-1 into one type for comparison with the “Grassland” of NLRS data. Especially in Northwest China (Figure R7) where there is a large-scale mixed distribution of “Grassland” and “Barren and sparse vegetation”, the strategy can be a suitable way to evaluate the overall performance of SinoLC-1.

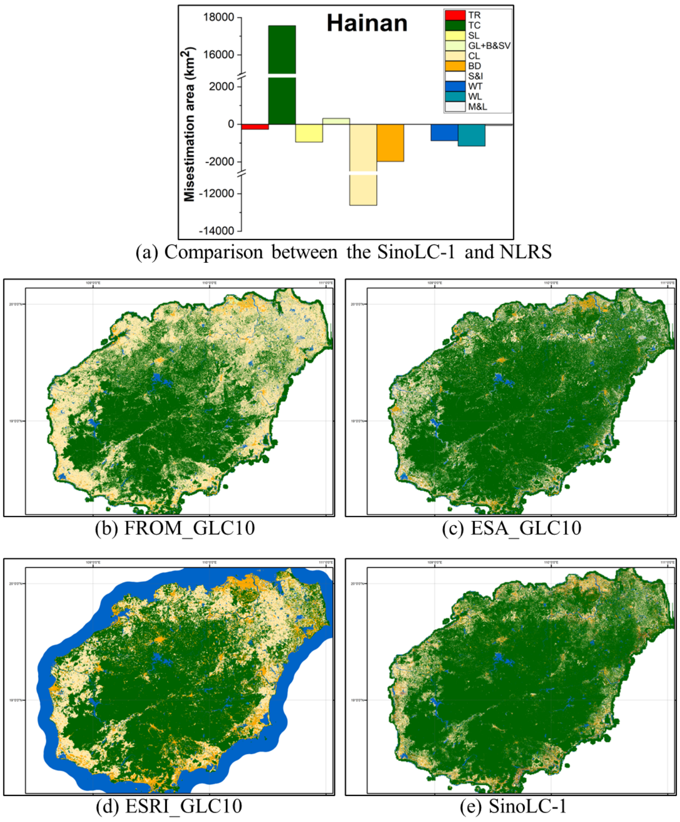

Figure R5. The qualitative and statistical results of the SionLC-1 in Hainan Province.

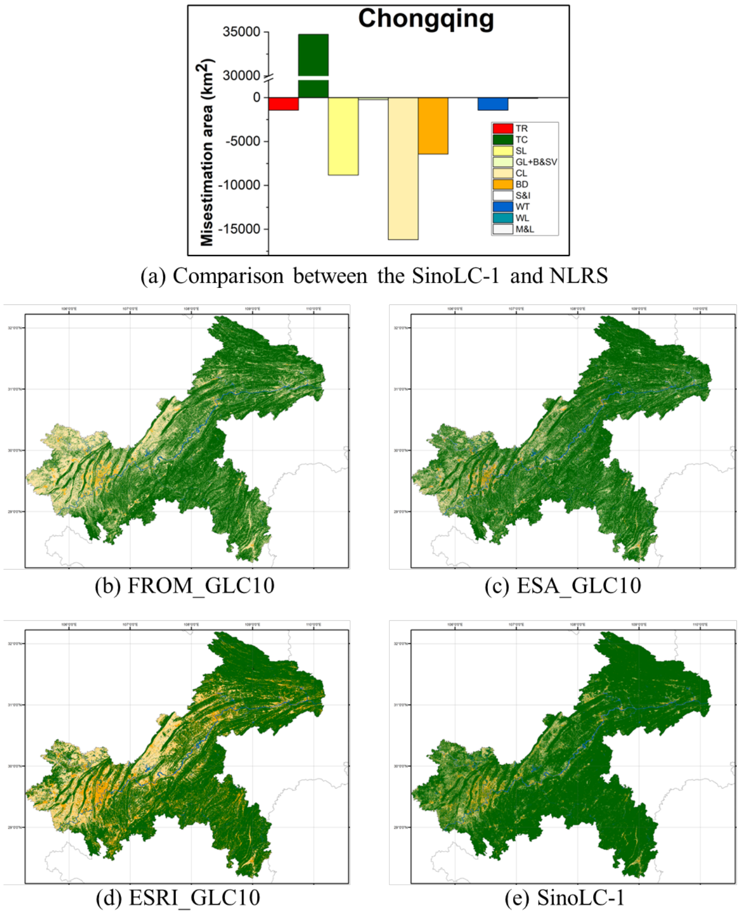

Figure R6. The qualitative and statistical results of the SionLC-1 in Chongqing Province.

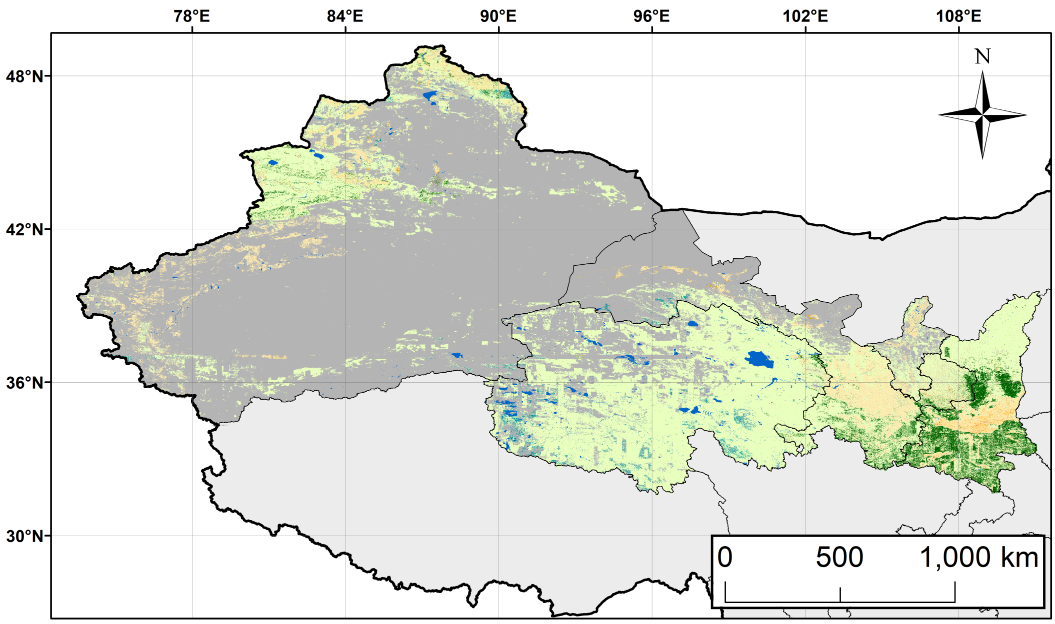

Figure R7. Demonstration of SinoLC-1 land-cover product in Northwest China.

Table R2. The definition of “Grassland” in the NLRS data.

Land-cover type of NLRS data	The definition of NLRS data
Grassland	1. Grasslands mainly composed of natural herbaceous plants, used for grazing or mowing, including grasslands where grazing prohibition measures are implemented. 2. A grassland artificially planted with grass. 3. Grassland improved by irrigation, drainage, fertilization, loosening, and replanting. 4. The surface layer of soil with a canopy density of less than 0.1. 5. Land used for scientific research, experimentation, and demonstration of herbaceous plants. 6. Due to factors such as engineering needs and improving the living environment, migration, and relocation have caused natural growth of herbaceous plants in villages, farmland, and other areas. 7. Land outside the scope of land acquisition for railways and highways, or the ditches without land acquisition, artificially planted and fixed for greening and beautifying the environment with herbaceous plants. 8. Land outside urban or villages which is artificially planted with herbaceous plants for greening and beautifying the environment. 9. Land used for cultivating herbaceous plant seeds. 10. In the grassland, supporting facilities that directly serve animal husbandry, such as land for storing forage and feed, drinking water for humans and animals, medicinal baths, shearing points, fire prevention, etc 11. Due to natural disasters causing damage to the arable land, naturally growing herbaceous land which is difficult to immediately restore cultivated through simple reclamation measures. 12. The mixed growth of herbaceous plants, trees, and shrubs cannot be distinguished, and the land is mainly composed of herbaceous plants. (Among them, the tree canopy density is less than 0.1 and the shrub coverage is less than 40%). 13. Land for planting commercial turf.

(7) The newly developed high-resolution land cover data will be very useful in land planning and management. It can also be used as the reference data for land cover classification at a relatively coarse resolution. Thus, it is important to know the acquisition time. Figure 24 shows the image capture time and the number of image tiles in different years, but I did not find the acquisition time file on the zenodo.

Response:

Thanks for the comment. We agree that image capture time is very important to land planning and management when users utilize the proposed land-cover product. We have supplemented the acquisition time file (organized as a tiff file) on the Zenodo website (https://zenodo.org/record/8214467) and added other materials to improve the data availability and integrity. Besides, according to the data update, we also released the new version of "User Guide v2.4" on the Zenodo website (https://zenodo.org/record/8214871).