Dear Topical Editor and Reviewer:

On behalf of my co-authors, we thank you very much for reviewing our manuscript and giving us a lot of useful comments and suggestions. We appreciate the comments on our manuscript entitled “GLC_FCS30D: The first global 30-m land-cover dynamic monitoring product with a fine classification system from 1985 to 2022 using dense time-series Landsat imagery and continuous change-detection method” (essd-2023-320).

We have revised the manuscript carefully according to the comments. All the changes were high-lighted (red color) in the manuscript. And the point-by-point response to the comments of the reviewers is also listed below.

Looking forward to hearing from you soon.

Best regards,

Prof. Liangyun Liu

liuly@radi.ac.cn

Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences

No.9 Dengzhuang South Road, Haidian District, Beijing 100094, China

Comments for essd-2023-320

The manuscript focused on the effort of mapping global 30-m land cover and change in 1985-2022. An advanced continuous change detection and classification approach was chosen with Landsat imagery to characterize global 35 different land cover types. Also, independent reference dataset was collected to validation the mapping product. Other third-party validation datasets were also implemented to provide additional validations. Relatively high mapping accuracy was achieved from the mapping products. The basic idea and approach are interesting, and the data presented by the authors have good potential for an interesting paper. However, there are gaps and lack of clarity in some parts of the manuscript, in particular method section, followed by the introduction and result.

Great thanks for the comments. The manuscript has been improved according to your comments and suggestions.

Major comments:

1. Using data from stable areas as training dataset is a good way. However, this study used the mapped data itself in one year as the training source. It is not clear for me is that the stable area is just from 2020 detection or from all year detection. Or in other way, how these stable areas were determined need additional clarification.

Great thanks for the comment. The quality of training samples is the key in land-cover mapping and change monitoring. In this study, all training samples are derived from temporal stable areas during 1985-2022 instead of the single year of 2020. The stable areas were determined by the continuous change detection (CCD) algorithm in Section 3.2. More clarification was added to describe how the training samples were derived. The manuscript has been strengthened in Section 3.3.1 as:

"we combined the GLC_FCS30-2020 prior dataset and **the change-detection mask (derived from the CCD algorithm described in Section 3.2)** to obtain the spatiotemporally stable training samples. Specifically, temporally stable areas are known to have higher mapping accuracy (Yang and Huang, 2021; Zhang and Roy, 2017; Zhang et al., 2023); thus, **we first used the aforementioned CCD mask to retain these temporally stable areas during 1985-2022, and then overlap them into the GLC_FCS30-2020 maps to determine their land-cover labels**. Next, Radoux et al. (2014) emphasized that land-cover transition areas usually were subject to more serious misclassification problems and that the homogeneous land-cover pixels had a higher probability of achieving acceptable accuracy. Therefore, we used the morphological erosion filter of 3 pixels × 3 pixels to refine these temporally stable areas into spatiotemporally homogeneous areas."

2. The data in the stable areas do not guarantee the land cover types in these area are correct. More discussions are needed to justify the rational of this training data selection.

Yes, we agree that the training samples derived from the stable areas cannot be guaranteed to be completely correct. In this study, except for the temporally stable constraint, the spatial homogeneity checking (the morphological erosion filter of 3 pixels × 3 pixels) also applied to optimize the training samples. Next, the GLC_FCS30 in 2020 was demonstrated to achieve an overall accuracy of 82.5%, and showed obvious advantages in mapping accuracy and diversity of land-cover types comparing with other land-cover products. In addition, our previous studies in generating the GLC_FCS30 land-cover maps have demonstrated that the local adaptive random forest classification models also showed great robustness to the erroneous training samples as Figure S1 (Zhang et al., 2021):

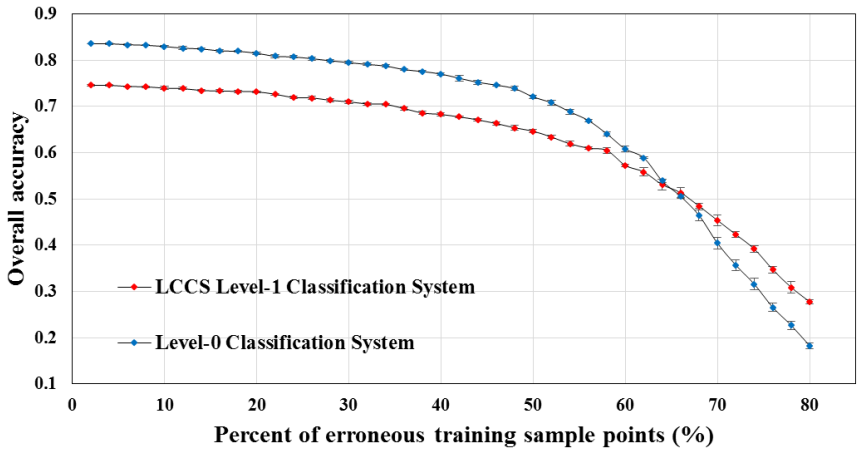


Figure S1. Sensitivity analysis showing the relations between the overall classification accuracy at two different classification system and the percentage of total samples and erroneous sample points. The Figure S1 came from the work of Zhang et al. (2021).

The quantitative relationship in Figure S1 indicated that overall accuracy of two classification systems (level-0 and LCCS level-1) generally decreased with the increasing of percentage of erroneous sample points. It remained relatively stable when the percentage of erroneous training sample was controlled within 30%, and decreased obviously after exceeding the threshold of 30%.

Zhang, X., Liu, L., Chen, X., Gao, Y., Xie, S., and Mi, J.: GLC_FCS30: global land-cover product with fine classification system at 30 m using time-series Landsat imagery, Earth Syst. Sci. Data, 13, 2753-2776, https://doi.org/10.5194/essd-13-2753-2021, 2021.

In the revised manuscript, the discussion about the quality of derived training samples have been added in the training sample Section as:
Benefiting from the temporally stable checking during 1985-2022, spatial homogeneity analyzing, and the overall accuracy of 82.5% in GLC_FCS30-2020 products, these spatiotemporally stable areas are retained to generate the training samples. It should be noted that the spatiotemporally stable areas are not guaranteed to be completely accurate, that is, a small number of derived training samples might be mis-labeled. Fortunately, previous studies in large-area land-cover mapping demonstrated that the random forest classification models (adopted by this study in Section 3.3.2) is highly robust to the erroneous training samples (Gong et al., 2019b; Mellor et al., 2015; Zhang et al., 2021b). For example, Gong et al. (2019b) found that the overall accuracy kept relatively stable when the proportion of erroneous training samples is controlled within 20%. Thus, the used spatiotemporally stable areas can be supported to derive confident training samples and further ensure the quality of land-cover dynamic monitoring.

3. A 26-time steps were used to map land cover by every five years before 2000 and annually after 2000 (Page 3, lines 119-120). However, why such steps are used did not given. The authors need to explain because you may mapped annual land cover and change in these periods without having too much extra cost.

Great thanks for the comment. Yes, we agree that we can generate the annual land-cover change maps during before 2000 without too much extra cost. However, why we still choose to update land-cover changes with 5-years interval before 2000 are the Landsat observations before 1999 only came from the single Landsat 5 TM sensor, which meant the valid observation imagery are too sparse. As we all known, the number of observation

data greatly affected the land-cover mapping and change detection, thus, we **have to sacrifice the temporal cycle (from one year to 5-years) to improve the land-cover change monitoring quality**. In this study, we combine the satellite observations from two years before and after the nominal center year from 1985 to 1995; for example, we update the land-cover maps in 1995 using all available imagery from 1993 to 1997.

The reasons why we choose the GLC_FCS30D updated every 5 years before 2000 has been added in the Introduction Section as:

In this study, we had the following three aims: 1) use the continuous change-detection algorithm and full time-series Landsat observations to generate the first global 30-m land-cover dynamic products with fine classification system (GLC_FCS30D) from 1985 to 2022, which contains 35 fine land-cover subcategories with 26 time-steps (maps updating every five years before 2000 and annually after 2000). **It should be noted that the GLC_FCS30D updated every five-years before 2000 due to the sparse availability of Landsat 5 imagery, thus, we combine the satellite observations from two years before and after the nominal center year from 1985 to 1995 for ensuring the mapping accuracy of GLC_FCS30D before 2000.**

In the Method Section, the reason is also been added as:

Last, since the land-cover distribution was usually related to the topographical environment, for example, croplands and water bodies are mainly distributed in flat areas, three topographical variables (elevation, slope, and aspect), calculated from a global 30 m DEM dataset (named as: ASTER_GDEM) (Tachikawa et al., 2011), were also imported. **In addition, due to the limited storage capacity and satellite–ground data-transmission capacity of early satellites, the density of Landsat imagery is sparse before 2000 (only Landsat 5 single-satellite acquired data) (Roy et al., 2014b). We choose the coarse temporal cycle of 5-years for ensuring the mapping accuracy before 2000, that is, the satellite observations from two years before and after was used for the nominal center year. For example, we update the land-cover maps in 1995 using all available imagery from 1993 to 1997.** In total, there were 49 multisource features, including 40 phenological spectra features, 6 texture features, and 3 topographical variables.

Meanwhile, the Section 4.5 also explained that one of our further works would combine multisourced remote sensing imagery to achieve the goal of global annual land-cover change monitoring before 1985 as:

Additionally, due to the limited storage capacity and satellite–ground data-transmission capacity of early satellites, the density of Landsat imagery is sparse before 2000 (only Landsat 5 single-satellite acquired data) (Roy et al., 2014b). In this study, we combine the satellite observations from two years before and after the nominal center year from 1985 to 1995; for example, we update the land-cover maps in 1995 using all available imagery from 1993 to 1997. However, a previous study found that northeastern Asia did not have any valid Landsat observations before 2000 (Zhang et al., 2022), which means some land-cover changes could not be captured in the GLC_FCS30D in these areas before 2000. To solve the problem of missing and sparse observations, a useful solution is to fuse multisourced remote-sensing imagery. For example, Zhang et al. (2021c) combined Landsat and Sentinel-2 imagery to track tropical forest disturbances with overall accuracy of more than 87%. Therefore, further work will investigate the feasibility of integrating Sentinel 1/2, SPOT, MODIS, and AVHRR imagery as auxiliary datasets to achieve the annual land-cover monitoring before 2000 and further ensure the land-cover monitoring quality.

The Section 5 “Data availability” also added the description about the different updating cycle as:

The developed GLC_FCS30D dataset can be freely accessible via https://doi.org/10.5281/zenodo.8239305 (Liu et al., 2023). To allow users to better select this dataset, it is saved as 961 5° × 5° independent tiles. Each tile is named as ‘GLC_FCS30D_yyyyYYYY_E/W**N/S**.tif’, in which ‘E/W**N/S**’ represents the longitude and latitude coordinates of the top-left corner, and yyyy and YYYY are the start and end years of the land-cover

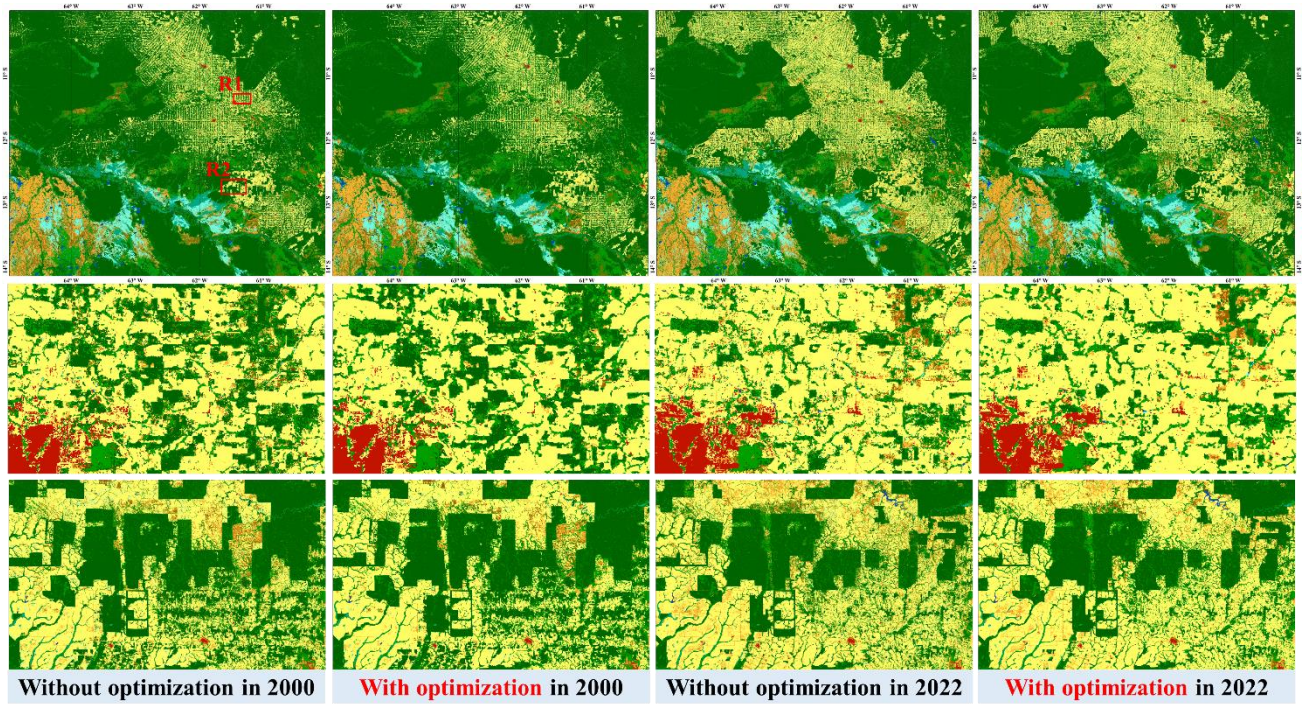
change monitoring. The GLC_FCS30D contains 26 time-step maps from 1985 to 2022, updated every five years before 2000 and annually from 2000 to 2022. **It should be noted that the GLC_FCS30D adopted the 5-years cycle before 2000 because of the sparse availability of Landsat 5 imagery at early stage, thus, we sacrificed temporal cycle in guaranteeing the land-cover mapping accuracy.**

4. A moving window of 3x3x3 are used to optimize temperature consistency of mapped land cover. However, such smooth window approach could also remove some rare classes that may not be larger than these moving windows. Some sensitivity analysis should be performed to compare the classification with and without moving window.

Great thanks for the comment. Yes, the temporal consistency optimization using the moving window of 3x3x3 maybe smooth some rare classes. In most case, the rare land-cover types, which are smaller than the windows, still can be retained because of the empirical of 0.5. It should be noted that the method was firstly designed for optimizing the impervious surfaces in the work of Li et al.,(2014), (the impervious surfaces are the representative rare land-cover type), and their analysis showed that the optimization method still retained most small rural impervious surfaces.

Li, X., Gong, P., and Liang, L.: A 30-year (1984–2013) record of annual urban dynamics of Beijing City derived from Landsat data, Remote Sensing of Environment, 166, 78-90, https://doi.org/10.1016/j.rse.2015.06.007, 2015.

In this revision, we also analyzed performance of the optimization method at two typical tiles (Amazon's deforestation and urban expansion in Yangtze River Delta) with or without the moving window as:



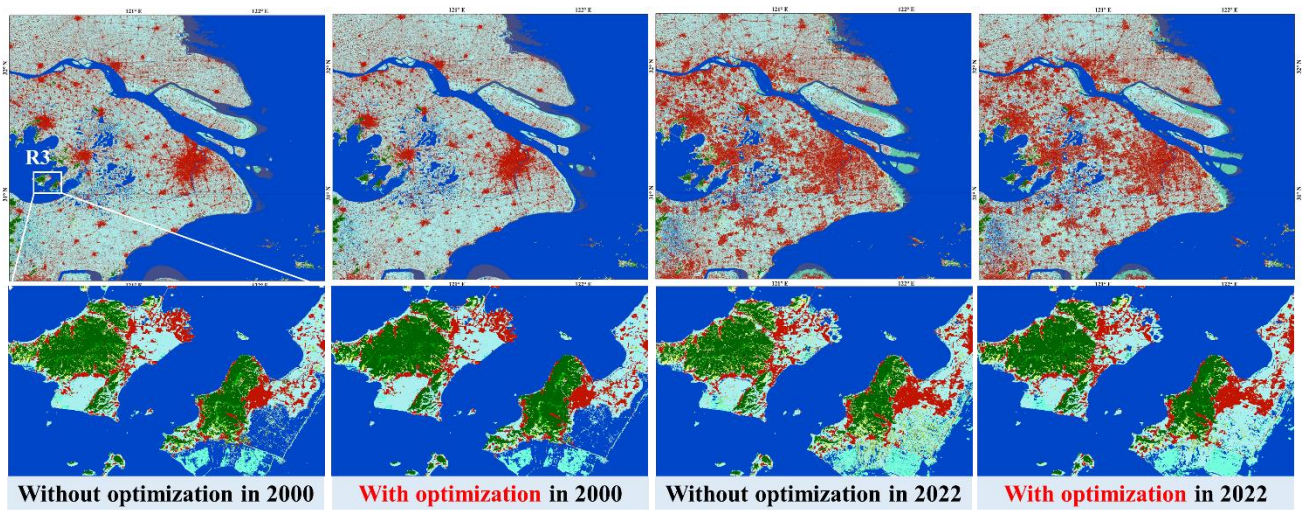


Figure S2. The comparisons before and after temporal consistency optimization in the Amazon's deforestation areas and urban rapid expansion area and their randomly enlargements.

Intuitively, the temporal consistency optimization preserved the spatial details well, that is, most small or fragmented forests and impervious surfaces were retained completely. Meanwhile, this post-processing showed great ability to deal with the 'salt and pepper' noise caused by the pixel-based classification, especially in the R1 and R3, namely, the optimized maps were 'visually-clear' than the origin maps

In summary, the comparisons indicates that the temporal optimization still retain the rare land-cover types and further improve the land-cover temporal consistency by removing some 'salt and pepper' noise caused by the pixel-based classification. It should be noted that this temporal-consistency optimization algorithm attached wide attentions in land-cover post-processing, for example, Yang et al., (2021) applied the same method to optimize the China annual land-cover products (CLCD) during 1990-2019, and Xie et al. (2021) used the similar method to improve the quality of land-cover change monitoring in Beijing during 2000-2021.

Yang, J. and Huang, X.: The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019, Earth Syst. Sci. Data, 13, 3907-3925, https://doi.org/10.5194/essd-13-3907-2021, 2021.

Xie, S., Liu, L., Zhang, X., and Yang, J.: Mapping the annual dynamics of land cover in Beijing from 2001 to 2020 using Landsat dense time series stack, ISPRS Journal of Photogrammetry and Remote Sensing, 185, 201-218, https://doi.org/10.1016/j.isprsjprs.2022.01.014, 2022.

In the manuscript, the discussion about the temporal consistency optimization was also added in the Section 4.5 as:

To ensure stability of the GLC_FCS30D, a spatiotemporal consistency optimization algorithm, which was widely used in impervious surface change optimizations (Li et al., 2015; Zhang et al., 2022), was applied. It makes full use of the spatiotemporal neighbor pixels to calculate the land-cover homogeneity, and then remove the 'salt and pepper' noise caused by the pixel-based classifications. The qualitative comparisons in Amazon's deforestation areas and China's urban expansion areas (Figure S2 in the supplement material) also showed that the spatiotemporal consistency optimization can improve the data quality of GLC_FCS30D by suppressing 'salt and pepper' noise and optimizing the temporal consistency. Similarly, Yang and Huang (2021) used this algorithm to optimize China's annual land-cover products during 1999 during 1990-2019, and found that it improved the mapping accuracy of the time-series land-cover dataset.

5. Figure 5 (b) and Figure 6 explain the global land cover change and changes in different regions. A global land cover spatial distribution in 1985 should be included so that reader can see the difference from 2022 map listed in Figure 4.

Great thanks for the suggestion. The GLC_FCS30D in 1985 has been added into the Supplement material:

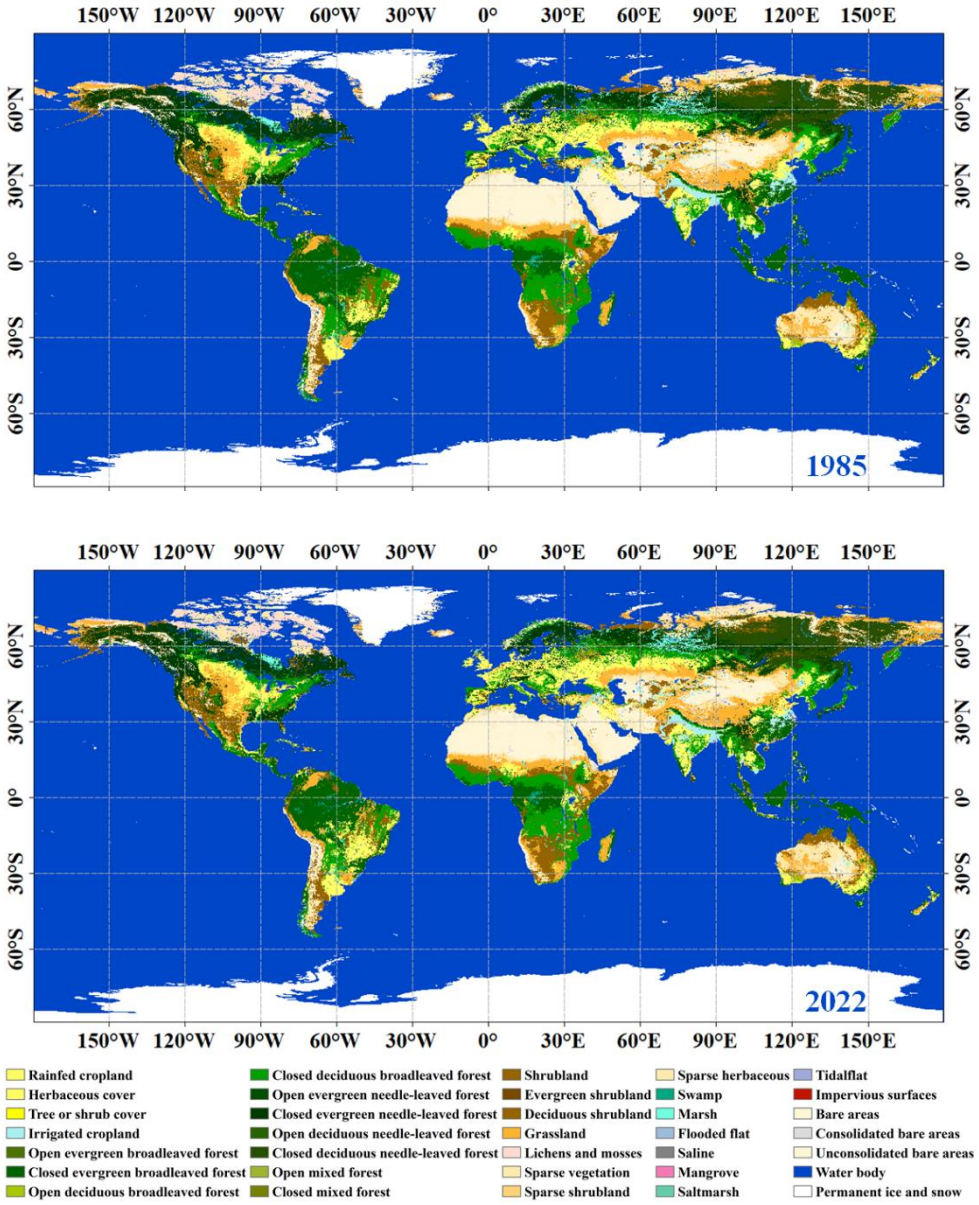


Figure S3. The overview of the GLC_FCS30D land-cover maps in 1985 and 2022.

Detail comments:

1. Page 1, line 21. "GLC-FCS30D is described as the first global 30-m land-cover dynamic monitoring dataset," The statement is not true. Several global 30-m land cover data, Copernicus Global land Service (100 m) 2015-2019, China Globeland30 (2000 and 2010 in 30-m), 9-class annual land use and land cover by Impact Observatory, are also available.

Great thanks for your comment. The statement has been revised as "GLC_FCS30D is described as the novel global 30-m fine land-cover dynamic monitoring dataset".

The similar statement on Page 2, line 75-79 needs to be changed to include the current existed global land cover products.

Based on your suggestion, the CGLC100 and GlobeLand30 land-cover products have been added into the Introduction Section.

In the early stage, GLCCM mainly relied on the time-series MODIS, AVHRR, and Project for Onboard Autonomy (PROBA)-V imagery (**Buchhorn et al., 2020**; Friedl et al., 2010); for example, Sulla-Menashe et al. (2019) generated a global 500-m annual land-cover products (MCD12Q1) from 2001 to present using time-series MODIS imagery with an overall accuracy of 73.6%.

Recently, benefitting from the free access to fine-resolution satellite imagery and powerful computing and storage capabilities, especially after the rise of cloud computing [such as Google Earth Engine (Gorelick et al., 2017) and Microsoft Planetary Computer], fine-resolution land-cover dynamic monitoring is experiencing rapid development. Correspondingly, numerous national and global 30-m land-cover dynamic products have been developed (**Chen et al., 2015**; Homer et al., 2020; Liu et al., 2021a; Potapov et al., 2022; Yang and Huang, 2021; Zhang et al., 2022).

Buchhorn, M., Lesiv, M., Tsendbazar, N.-E., Herold, M., Bertels, L., and Smets, B.: Copernicus Global Land Cover Layers—Collection 2, Remote Sensing, 12, 1044, https://doi.org/10.3390/rs12061044, 2020.

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., and Mills, J.: Global land cover mapping at 30m resolution: A POK-based operational approach, ISPRS Journal of Photogrammetry and Remote Sensing, 103, 7-27, https://doi.org/10.1016/j.isprsjprs.2014.09.002, 2015.

2. Page 4, lines 135-136. Why only listed the spectral difference between ETM+ and OLI?

Thanks for the comment. The previous studies have explained that the TM and ETM+ shared the same spectral response function, and the spectral difference between TM and ETM+ are directly ignored (Roy et al. 2016). The make the sentence clearer, it has been revised as:

Then, although the Landsat 5, 7, 8, and 9 missions share similar spectral bands, the wavelength differences between the **TM,** ETM+ and OLI cannot be ignored. Relative radiometric normalization was applied to the TM and ETM+ imagery using the transformation coefficients suggested by Roy et al. (2016).

3. Page 11, lines 307-308/ sample size of 8000 and 600 are recommended. Are these sample thresholds for one tile?

Thanks for the comment. Yes, the recommended sample size of 8000 and 600 are suitable for all 961 5° × 5° geographical tiles. The corresponding explanation has been revised as:

Thus, the maximum and minimum sample size for abundant and rare land-cover types were suggested as **8000 and 600**, came from the study in Zhu et al. (2016), for avoiding the extremes of sample sizes. Next, the GLC_FCS30-2020 products were split into 961 5° × 5° geographical tiles, and we used the areal-proportional sampling strategy and **two sample balancing thresholds (8000 and 600 for maximum and minimum sample size)** to allocate the training samples from the spatiotemporally stable areas **in each 5° × 5° geographical tile**. Last, the impervious surface and wetland samples were excluded because both have been independently developed as the thematic datasets in Section 2.3 and 2.4.

4. Page 14, lines 400 -401. What are 10 major land cover types?

Great thanks for the comment. The 10 major land-cover types have been explained in the Table 1, which include **cropland, forest, shrubland, grassland, tundra, wetland, impervious surface, bare areas, water body,**

permanent snow and ice. To make the sentence clearer, it has been revised as:

The GLC_FCS30D adopts a fine classification system containing 35 subcategories, for which we applied an analysis protocol into the basic classification system and the LCCS level-1 validation system, **whose details were explained in the Table 1 and contained 10 major land-cover types and 17 fine land-cover types (Table 1)**, respectively.

5. Figure 5a. The graphic is good to show the spatial distributions of land cover change intensity. However, there are no further discussions about these changes, especially for some area experiencing very change high intensity, e.g., changes in the southern Hudson Bay area in Canada, western coast of African, and eastern side of Australia.

Great thanks for the comment. Actually, these areas experienced very change high intensity have been independently discussed as:

"Obviously, global land-cover has experienced significant changes over the past 37 years, mainly in the following three typical areas: 1) tropical rainforest peripheral areas in South America and Southeast Asia, in which deforestation is the dominant cause; 2) wetland and water-body intermingling areas, such as **North America** and northern Asia, i**n which water bodies and wetland were transformed into one another due to different annual water levels. In the GLC_FCS30D the water body land-cover type represents permanent water during the year (it may be wetland in other years)**. 3) The semi-arid areas in **Australia**, Central Asia, and **western Africa**, where land cover (such as sparse vegetation or bare land) is directly affected by precipitation and temperature. For example, if there is sufficient precipitation in the year, the sparse vegetation and some bare land would be covered by grass in semi-arid areas. Similarly, the work of Winkler et al. (2021) revealed that these semi-arid areas experienced serious and frequent land-cover changes."

Namely, the changes in the southern Hudson Bay area in Canada are obvious because the water bodies and wetland were transformed into one another due to different annual water levels. As for the western coast of African and eastern side of Australia, both of them belongs to the semi-arid areas, where land cover (such as sparse vegetation or bare land) is directly affected by precipitation and temperature. For example, if there is sufficient precipitation in the year, the sparse vegetation and some bare land would be covered by grass in semi-arid areas.

6. Full land cover names for these synonyms names in Table 2 & 3 need to be listed.

Great thanks for the comment. The full land-cover names of these synonyms names have been moved into the Table 1 based on the Reviewer 2 suggestions in the first round. In this revised manuscript, the notes about the synonyms have been added in the **lower left corner of Table.**

7. Figure 10. What happened for the UA of bare land in 2005-2015? What caused such sudden increase in UA?

Great thanks for the comment. The sudden increase of U.A. in Figure 10 is Ice and snow (instead of bare land). The cause is that the Ice and snow is a greatly sparse land-cover type in the U.S., its validation sample size in the LCMAP is also small and a little misclassified grass/bare land pixels are correctly identified as snow and ice during 2005-2014. The manuscript has been revised as:

The large fluctuations in U.A. for ice and snow are attributed to: 1) the small size of ice and snow samples in the LCMAP_Val dataset, 2) a few misclassified grass/bare land pixels correctly identified as snow and ice during 2005-2014.