Manuscript essd-2022-224, Second Revision
Submitted to: Earth System Science Data

CALC-2020: a new baseline land cover map at 10 m resolution for the circumpolar Arctic

Chong Liu, Xiaoqing Xu, Xuejie Feng, Xiao Cheng,
Caixia Liu, and Huabing Huang

Responses to Referee #2

December 15 2022

General comments

The manuscript has been improved according to the comments. However, I still have three concerns about the revised manuscript:

Response: We would like to thank the reviewer for the encouragement, and pointing out issues that help improve the research and the quality of this paper. According to these comments, we further revised the manuscript and added figures/tables to ensure it meets the quality and high standard of ESSD. Below we provided our detailed responses to the points raised.

Major comments

1. The novelty of the method and Dataset is a little weak;

Response: We agree that the core of any research lies in its novelty and contributions to the scientific community. With the help from two reviewers and the topical editor, we respectively argue that our revised manuscript has novelty from the following two aspects.

1) We proposed a new paradigm for fine resolution Arctic land cover mapping using remotely sensed big data streams. As stated in our paper, the terrestrial Arctic environment is a fundamentally different ecosystem from those at lower latitudes, which calls for reconsideration of every aspect of satellite remote sensing, including data acquisition, land cover legend design, classification, and accuracy assessment. To achieve this goal, we developed a framework that incorporates data from multiple satellites observing the Arctic, and migrated sample from pre-existing knowledge. We modified the FROM-GLC class scheme to represent the biome diversity across the terrestrial Arctic, and used locally adaptive (country-specific) RFC models to reduce the unfavorable impacts incurred by geographical variability. To confirm the reliability of the proposed framework, the classification result was evaluated against validation sample data, in-situ records, and contemporary land cover products.

2) More importantly, the developed CALC-2020 dataset can contribute to the scientific community by providing the first spatially continuous map of circumpolar Arctic land cover at 10 m resolution. Based on this map, new insights can be given into Pan-Arctic land cover composition and distribution that have not been fully documented in existing publications or datasets, and therefore advances our understanding of a continuously changing Arctic as well as its global environmental impacts.

Based on the above two aspects, we kindly hope the reviewer re-evaluate our work after its 2rd-round revision.

the validation accuracy of the CALC is still low (shrub tundra, which is the advantage of the dataset).

Response: We gave critical thinking and double-check the sample-based validation procedure of CALC-2020. Some key points are clarified as follows.

1) Following reviewer#1's suggestion, we adopted "good practices" by Olofsson et al. (2014), and built the error matrix as well as associated sample-based accuracy statistics of the CALC-2020 map. The OA of the CALC-2020 land cover product for the circumpolar Arctic is 79.3±1.0% (95% confidence interval), which is reasonable and comparable to most regional-scale Arctic land cover mapping efforts (typically ranging from 65% to 85%).

2) As pointed out by the reviewer, the stratified error-adjusted PA estimate of shrub tundra is 47.1±5.7%, due primarily to the influence of estimation weights (area proportions of map classes). Here we refer to a recently published, continental-scale land cover product using the same "sample-adjusted" validation approach (GLanCE for North America by Friedl et al., 2022), and found a similar PA for the shrub biome (40.8 ± 6.0). It should be noted that all accuracy metrics reported in the manuscript are based on error matrix of area proportion (Olofsson et al., 2014), rather than traditional confusion matrix of sample counts. As shown in **Table R1**, the traditional confusion matrix-derived PA of shrub tundra (SRT) is 68.1%, which is much higher than its stratified error-adjusted estimation. To avoid possible misunderstanding, we added **Table R1** as **Table S3** in **Supplement**, and clarified this point in the revised manuscript:

Table R1 Accuracy statistics of the CALC-2020 map based on traditional confusion matrix of sample counts. Abbreviations of land cover are given in the main text.

Class	CRO	FST	GRT	**SRT**	WET	OWT	LAM	MMI	BAR	IAS
UA (%)	93.5	75.0	81.9	**84.7**	81.0	93.7	76.8	95.9	75.1	74.0
PA (%)	100.0	100.0	94.3	**68.1**	79.6	63.7	78.1	100.0	63.0	90.1
OA (%)						79.6				
Kappa						0.735				

Page 11, **Lines 261-264 in the revised manuscript**: It should be noted that all metrics reported in **Table 3** are based on error matrix of area proportion, therefore inevitably different with those derived from traditional confusion matrix of sample counts. For example, the traditional confusion matrix-derived PA of shrub tundra is 68.1% (**Table S3**), whereas its stratified error-adjusted PA estimate is lower, due primarily to the influence of estimation weights (area proportions of map classes).

3) We did recognize the challenge of shrub tundra mapping by the current version of CALC-2020, although it was developed by incorporating Optical, SAR and DEM images. Some researches (e.g., Gong et al, 2016) suggested improved classification performance of shrubland by using the vegetation height as an auxiliary feature metric. However, such a geospatial dataset is currently unavailable for the entire terrestrial Arctic. As a result, we expect a higher, more consistent accuracy level of CALC as new multi-source Earth observation data become available in the future. We clearly highlighted this issue in the **Discussion** section of the revised manuscript:

Page 20, **Lines 417-420 in the revised manuscript**: With updates of more multi-source Earth observation data in the future (e.g., vegetation height), the development of Version 2 CALC product will become a possible topic, which is expected to have a hierarchical classification scheme and improved mapping performances for some specific biomes (e.g., shrub tundra).

Reference 1: Olofsson, P., Foody, G. M., Herold, M., et al.: Good practices for estimating area and assessing accuracy of land change, Remote Sens. Environ., 148, 42–57, 2014.

Reference 2: Friedl, M. A., Woodcock, C. E., Olofsson, P., et al.: Medium Spatial Resolution Mapping of Global Land Cover and Land Cover Change Across Multiple Decades From Landsat, Front. Remote Sens., 3, 2022.

Reference 3: Gong, P., Yu, L., Li, C., et al.: A new research paradigm for global land cover mapping, Annals of GIS, 22, 87–102, 2016.

2. The confusion matrix cannot convince me espcially in the confusion between open water, ice and snow and barren land. In my opinion, the open water is easier to confused to the wetland.

Response: We have carefully double-checked our sample validation procedure to ensure each element in the error matrix (**Table 3**) correct. Given the generally large reflectance discrepancy between water and non-water covers, the less desirable performance of CALC-2020 in water extraction may seem unexpected. This highlights the distinctiveness of Arctic's geographical environment that can affect the spectral signal of water in space and time (Gong et al., 2016). At high latitudes, shallow water bodies are easily confused with barren lands because of the mixed pixel issue (**FigureR1a~b**). Moreover, the employed satellite images may only capture the freezing stage for some water pixels, which were misclassified as ice/snow in the CALC-2020 map (**Figure R1c**). We added **Figure R1** as **Figure S4** in **Supplement**, and explicitly discussed this issue in the revised manuscript.

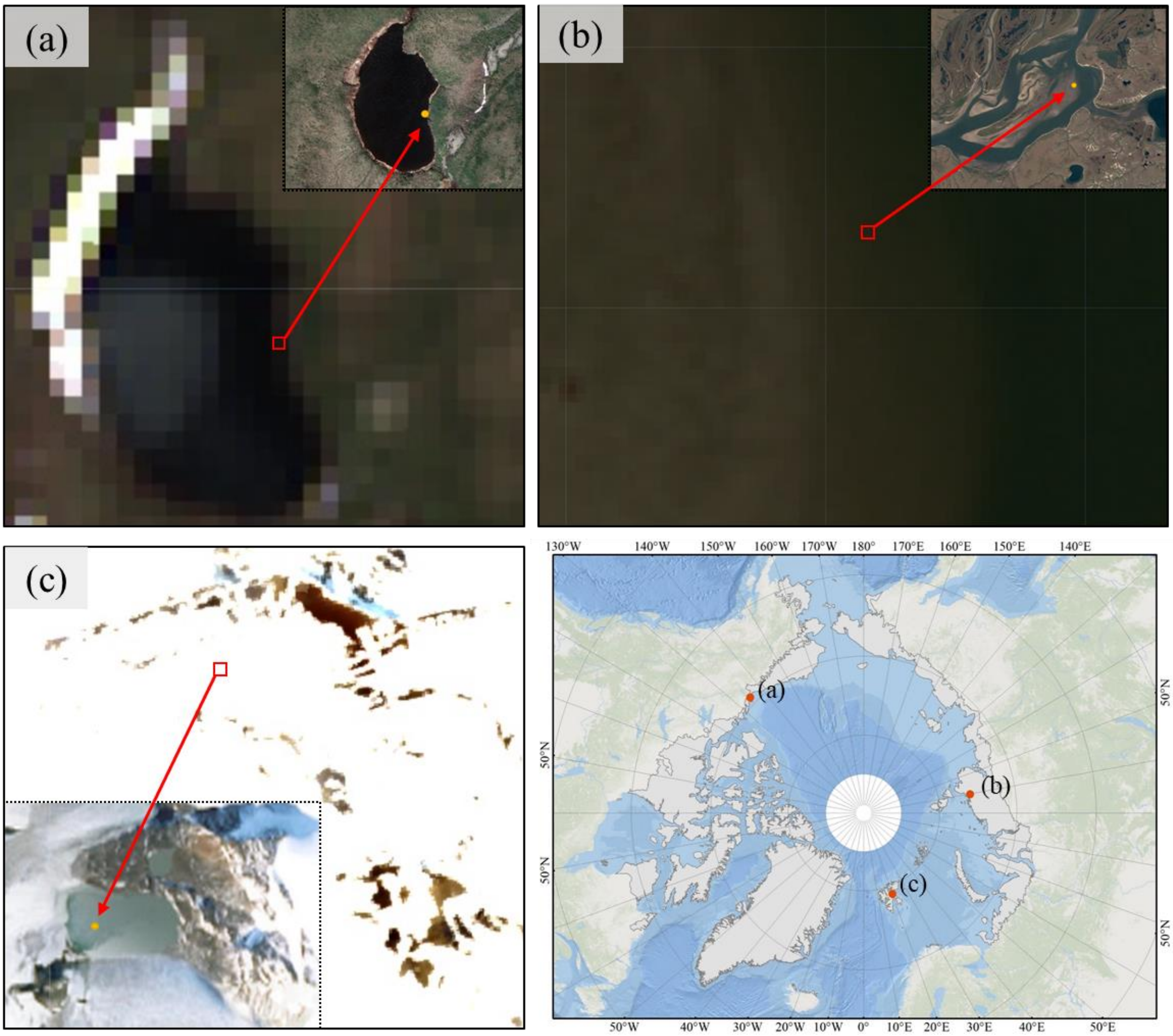

Figure R1 Sentinel-2 true-color images showing three typical examples of water sample points misclassified by CALC-2020. (a) Mixed pixel of lake water and barren land (centred at 69.1°N, 134.8°W). (b) Mixed pixel of river water and barren land (centred at 75.8°N, 99.7°E). (c) Frozen lake water confused with ice/snow (centred at 78.9°N, 20.1°E). The Google Earth VHR image is also displayed for each example.

Page 11, Lines 264-269 in the revised manuscript: Given the generally large reflectance discrepancy between water and non-water covers, the less desirable performance of CALC-2020 in water extraction may seem unexpected. This highlights the distinctiveness of Arctic's geographical environment that can affect the spectral signal of water in space and time (Gong et al., 2016). Specifically, shallow water bodies are easily confused with barren lands because of the mixed pixel issue (**Figure S4a~b**). Moreover, the employed satellite images may only capture the freezing stage for some water pixels, which were misclassified as ice/snow in the CALC-2020 map (**Figure S4c**).

Reference: Gong, P., Yu, L., Li, C., et al.: A new research paradigm for global land cover mapping, Annals of GIS, 22, 87–102, 2016.

3. The comparisons with other global products are unfair, I suggest the authors can provide the quantative metrics, and add some national land-cover products as the comparative datasets to demonstrate the advantages of the CALC-2020.

Response: We agree with the reviewer's suggestion. In the revised manuscript, we additionally compared our map with two national-scale land cover products, and calculated accuracy metrics of three global land cover datasets. The summary of major changes are as follows:

1) We compared the CALC-2020 map with NLCD (Jin et al., 2017) and Land Cover of Canada (Wulder et al., 2018) because they are currently the only two fine resolution national-scale land cover products covering parts of our study area. Both NLCD and Land Cover of Canada utilize hierarchical classification schemes, we therefore merged their level-2 class legends to that of CALC-2020 (level-1) for interpretation. It should be noted that our stratified validation sample was designed for the circumpolar Arctic, thus not suitable for quantifying the accuracy of national-scale products (Olofsson et al., 2014). As an alternative, we compared the land cover distribution of our estimation with those from NLCD and Land Cover of Canada at 30 m resolution. As shown in **Figure R2**, CALC-2020 map shows overall high agreements with both national-scale land cover products, reflecting its reasonable mapping performance in North America. We added **Figure R2** as **Figure S6** in **Supplement**, and explicitly discussed this issue in the revised manuscript.

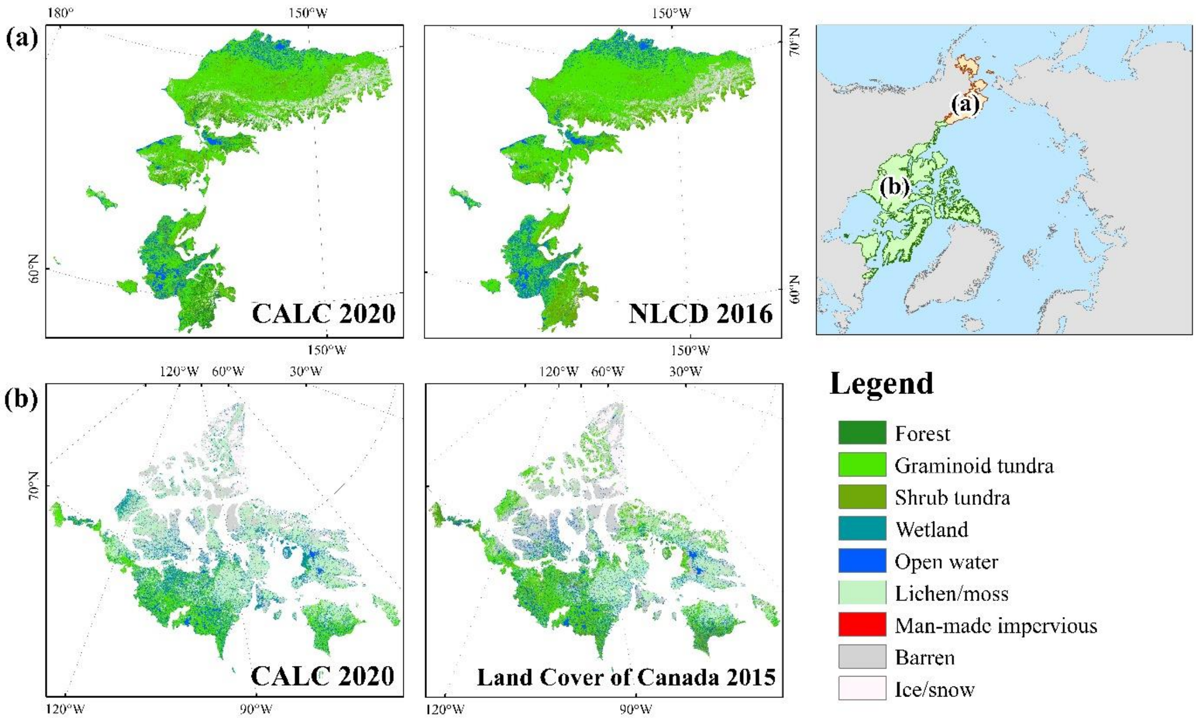

Figure R2 Comparison of CALC-2020 and two national-scale land cover products: NLCD 2016 (a) and Land Cover of Canada 2015 (b). The level-2 classification schemes of two national-scale land cover products are merged to that of CALC-2020 (level-1) for interpretation.

Page 15, **Lines 333-336 in the revised manuscript**: The reasonable performance of the CALC-2020 map for North America was also confirmed by referring to two national-scale land cover products: NLCD and Land Cover of Canada, both of which exhibit high agreement of land cover distribution pattern with CALC-2020 at their level-1 classification schemes (**Figure S6**).

2) We calculated OA, PA and UA of three global land cover products (ESA WorldCover, ESRI Global Land Cover, GlobeLand30) using the same sample-based evaluation approach applied to CALC-2020. To harmonize various classification legends to that of CALC-2020, the grass (ESA WorldCover, ESRI Global Land Cover) and wet tundra classes (GlobeLand30) were treated as equivalents of graminoid tundra and wetland, respectively. We reported limited classification accuracies of the three compared land cover products (**Figure R3**), with OAs ranging from 48.5% to 71.2%. In the meantime, three global-scale datasets exhibit wide PA and UA variations. The quantitative comparison result is generally consistent with a previous research (Liang et al., 2019), implying imbalanced performances across different Arctic land cover types mapped by pre-existing global-scale datasets. We added **Figure R3** as **Figure 7** in the revised manuscript, and offered relevant descriptions.

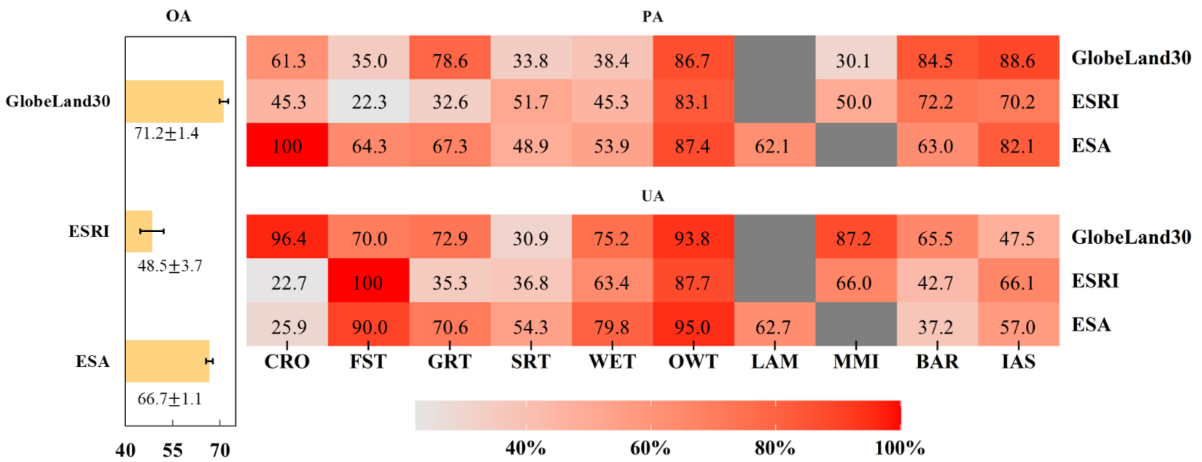

Figure R3 Accuracy statistics of three global land cover products for the circumpolar Arctic based on validation sample. Note that the grass (ESA WorldCover, ESRI Global Land Cover) and wet tundra classes (GlobeLand30) are treated as equivalents of graminoid tundra and wetland, respectively. The dark grey color represents the absence of specific class(es).

Page 10, **Lines 237-240 in the revised manuscript**: As an additional comparison to complement the inter-product evaluation, we used the validation sample shown in **Figure S3** to calculate accuracy metrics of three global land cover products. To harmonize various classification legends to that of CALC-2020, the grass (ESA WorldCover, ESRI Global Land Cover) and wet tundra classes (GlobeLand30) were treated as equivalents of graminoid tundra and wetland, respectively.

Page 14, **Lines 313-316 in the revised manuscript**: Using the same sample-based evaluation approach applied to CALC-2020, we reported limited classification accuracies of three global land cover products for the circumpolar Arctic (**Figure 7**), with OAs ranging from 48.5% to 71.2%. In the meantime, these global-scale datasets exhibit wide PA and UA variations, implying imbalanced mapping performances across different Arctic land cover types (Liang et al., 2019).

Reference 1: Jin, S., Yang, L., Zhu, Z., and Homer, C.: A land cover change detection and classification protocol for updating Alaska NLCD 2001 to 2011, Remote Sens. Environ., 195, 44–55, 2017.

Reference 2: Wulder, M. A., Coops, N. C., Roy, D. P., et al.: Land cover 2.0, Int. J. Remote Sens., 39, 4254–4284, 2018.

Reference 3: Olofsson, P., Foody, G. M., Herold, M., et al.: Good practices for estimating area and assessing accuracy of land change, Remote Sens. Environ., 148, 42–57, 2014.

Reference 4: Liang, L., Liu, Q., Liu, G., et al.: Accuracy Evaluation and Consistency Analysis of Four Global Land Cover Products in the Arctic Region, Remote Sens., 11, 2019.