Geocarto International

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tgei20

A simple, fast, and accurate method for land cover mapping in Mongolia

Thanh-Noi Phan, Batnyambuu Dashpurev, Felix Wiemer & Lukas W. Lehnert

GEOCARTO INTERNATIONAL
https://doi.org/10.1080/10106049.2022.2087759

A simple, fast, and accurate method for land cover mapping in Mongolia

Thanh-Noi Phan[a,b], Batnyambuu Dashpurev[a], Felix Wiemer[a] and Lukas W. Lehnert[a]

[a]Department of Geography, Ludwig-Maximilians-University Munich, Munich, Germany; [b]Department of Geodesy and Cartography, Faculty of Natural Resources and Environment, Vietnam National University of Agriculture, Hanoi, Viet Nam

ABSTRACT

Low accuracy of global land cover (LC) products at local and regional scales is concerning, because of its huge impact on downstream applications. Therefore, developing a method for high accuracy LC maps at local and regional scale is of utmost importance. Taking Mongolia as a case study, we proposed a simple, fast, and accurate method to produce annual LC maps with 250 m spatial resolution from 2001 to 2020 using MODIS data. Our products have higher spatial resolution and higher accuracy (Overall Accuracy ∼90%) compared to MCD12Q1. These new maps are critical e.g., for land degradation research, and desertification/forest cover monitoring. Especially, information on grassland ecosystems is of utmost importance for Mongolia since more than half of the country economically depends on grassland resources. Therefore, Mongolia will benefit from the new dataset. Furthermore, due to the simplicity of the method, it can be easily applied and transferred to other regions.

ARTICLE HISTORY
Received 4 April 2022
Accepted 3 June 2022

KEYWORDS
MODIS; land cover; large scale; semi-automatic; Google Earth Engine; Mongolia

1. Introduction

1.1. The importance of land cover data

Land use and land cover (LUC) information is considered one of the most important and fundamental pieces of information that represent the biophysical properties of the Earth's surface. This information is critical to understand changes on the Earth's surface, in the atmosphere, and the interaction between them. Furthermore, LUC data is widely applied and plays an important role in downstream applications. For example, the accuracy of the LUC information would directly affect the performance of ecosystem, hydrology, and atmosphere models since such models depend on LUC information as a key input (Gong et al. 2013; Zhang et al. 2021). That is why LUC receive high attention from scientific communities, policy makers, and international organizations (Ban et al. 2015; Zhang et al. 2021).

Due to the high demand of the LUC data and the availability of remote sensing data, remotely sensed data has been widely applied to produce land cover maps from local to

CONTACT Thanh-Noi Phan Phan.Noi@lmu.de

regional and globe scales at various spatial resolutions, ranging from 10 m to 1 km (Loveland et al. 2000; Friedl et al. 2010; Tateishi et al. 2014; ESRI 2020, CCI-2020). Among these, some fine spatial land cover products (i.e., 10 m to 30 m) are only available for specific years or specific regions. For example, the fine spatial resolution global land cover maps at 30 m (GlobeLand30) are only available for two years, 2000 and 2010, while the S2 prototype LC 20 m map of Africa 2016 is only available for the African Continent. Meanwhile, the time series land cover data (e.g., annual land cover) that are usually available at coarse spatial resolutions, i.e., from 100 m to 1 km, have been reported to lack sufficiently high quality for many downstream applications (Giri et al. 2005; Grekousis et al. 2015; Yang et al. 2017).

1.2. The accuracy of regional land cover maps is more important than ever

In a recent review on the effects of land use on climate, Pongratz et al. (2021) argued that LUC data is critically required to attribute climate change to human action and to understand the drivers of climate change. Consequently, the attention of the availability and accuracy of LUC products is increasing. From a global perspective, a large proportion of uncertainties arise from undetected or false detection of LUC changes (Friedlingstein et al. 2020).

In the recent literature, increasing numbers of studies investigate how the accuracy of LUC data might affect downstream applications (e.g., if LUC data is used as input for models etc.). For instance, according to Tulbure et al. (2021), one of the most significant problems with global land cover products is that many maps have not been validated following the recommendation of the best practices (Olofsson et al. 2014). Therefore, if global products are evaluated for specific regions, their accuracy is often lower as official reports for the global scale state (Bai et al. 2015). This has been verified in several studies across several regions. For example, Bai et al. (2015) used an independent reference dataset over China to validate five widely used global land cover products and found that some of these products had a very low accuracy. In their study, Zheng et al. (2021) compared the accuracy of five popular global land cover datasets at varying spatial resolutions (MCD12Q1, CGLS-LC100, FROM_GLC, GLC_FCS30 and GlobeLand30) in different years (2010, 2015, 2017, and 2020) over the coastal area of Tianjin City (China) and found that the accuracy of all products was less than official announcements. Liu et al. (2019) used 923 field sampling points over the Qiangtang Plateau in High Asia to evaluate the local spatial accuracy of seven popular global land cover datasets. They also found that local accuracy was lower compared to official announcements. Furthermore, the accuracy also varies between land cover classes as reported by studies evaluating specific land cover types. For instance, Pérez-Hoyos et al. (2017) compared global land cover products for cropland monitoring, while Majasalmi and Rautiainen (2021) evaluated the representation of tree cover in Finland in different global land cover products. Both studies revealed that target land cover classes varied substantially between land cover products with considerable over- or underestimation. It should be noted that the accuracy of tree cover in the global land cover datasets over Finland was very low, just ranging from 42 to 75%. Therefore, it is suggested that the accuracy reports of the global land cover products should not be used to infer conclusions about the applicability to specific regions (Tulbure et al. 2021).

However, due to the lack of local and regional land cover products with high thematic accuracy for any downstream application, many studies have used global land cover datasets for studies on regional or even local scales. For example, Siddiqui et al. (2021) used MODIS Land Cover datasets to define urban and non-urban areas to study urban heat islands in three cities in India. Ahmad et al. (2021) used MODIS land cover products as the main input to analyse the effect of climate and land use changes on stream flow of the Chitral river basin of Pakistan. Li et al. (2020) used MODIS land cover products to quantify the effects of land use/cover change on the Net Primary Productivity in China's terrestrial ecosystems from 2001 to 2012.

The emphasis had been on the "region matters" (i.e., global products have low accuracy at some regions) when using the global land cover datasets at regional scale in current times of global change. Tulbure et al. (2021) noted that in times when global change also has a strong regional impact, but at the same time enormous global (big) data and appropriate processing algorithms/techniques are available, it is of utmost importance to have global data sets that are regionally applicable. High quality open access data can help researchers and the policy makers to respond to and mitigate climate change.

1.3. Global land cover products in Mongolia

To have a closer look at the quality of the global land cover products over Mongolia, we performed two simple comparisons. The first comparison is between medium and finer spatial resolution products, MCD12Q1 500 m and GlobeLand30 (30 m). To compare, we resampled and rescaled these land cover products to 250 m with six land cover types. As shown in Figure A1, there are large differences between these two land cover products. The most pronounced disagreement is between grassland, bare land, shrub land, and forest areas, which are the dominant and important land cover types in Mongolia. Since local herders in Mongolia directly depend on grassland resources for their livelihoods, the accuracy of grassland information is of utmost important for the herder community (also for the Mongolian economy).

In the second comparison, we took new high spatial resolution global land cover products, the ESRI land cover 2020 and the ESA world cover 2020. Both products have 10 m spatial resolution, while exhibiting 10 and 11 land cover classes, respectively. For the comparison, we kept the original land cover classes as well as the spatial resolution (Figure A2).

Large differences were obvious between bare land, grassland, and shrub land; all of them are important land cover classes for Mongolia. Therefore, this inconsistency would affect any downstream application.

1.4. Regional land cover products and their role in global change studies

There is undoubtedly a high demand for regional land cover products at regional scales featuring a sufficiently high temporal resolution and coverage, e.g., at yearly interval over long time periods. For example, the annual land cover products at 250 m have been developed for global change, climate change, and global warming studies, however, this kind of data has only been available for developed regions such as Europe, Canada, and North America.

In recent years, several approaches have been developed to collect samples from existing land cover products to reduce costs for collecting reference data (Radoux et al. 2014; Wessels et al. 2016; Zhang and Roy 2017, Hermosilla et al. 2022). The advantage of these methods is that they enable one to collect samples automatically and quickly. Compared to manually selecting the samples, the automatic methods ensure homogeneous geographical/spatial distribution of samples, which is particularly useful for large study areas (i.e., at national or regional scales).

In addition, it should be mentioned that one of the most common error sources in collecting in-situ samples is that the selected point is not representative for the entire area covered by the satellite pixel. This is particularly true in complex landscapes and land cover types if working with medium or coarse resolution imagery. The resulting scaling issue affects the classification results (Elmes et al. 2020).

However, the biggest challenge of any automatic sample collection method is to extract the samples reliably and by minimizing effects arising from imbalanced data, spatial autocorrelation, and clustered distributions over the study area, which all happen often but are rarely studied. Figure A3 shows an example of these mentioned effects.

Finally, Mongolia is a large country (∼1,564,116 km^2) with a very low population density. As infrastructure is in bad condition and approximately 80% of road network is unpaved, travelling around within Mongolia is time-consuming. Therefore, using existing land cover products to collect samples is considered the best option to provide Mongolia with accurate time series of land cover data.

From the reasons and issues mentioned above, this study aims to develop a land cover classification method that is simple, fast, and accurate for all of Mongolia. Consequently, this study contributes to two major aspects:

i. It provides a simple, fast, and accurate classification method by using open source resources (MODIS images, digital elevation data, global land cover products, and Random Forest classifier) and powerful cloud computing (Google Earth Engine). It is applicable for a wide range of users from the scientific community who may be less familiar with remote sensing and/or classification techniques, as well as governments and international organization, who want to apply remote sensing land cover classification over other study areas. Furthermore, due to the simplicity of the method, it can be applied and transferred to other regions easily.
ii. A yearly open access land cover map for entire Mongolia from 2001 to 2020 is made available, which allows researchers to directly apply the data to subsequent applications. Amongst others, researching vegetation condition, vegetation change, or land degradation monitoring over Mongolia for the last twenty years are possibilities.

2. Data and method

2.1. Data used

To produce land cover maps of Mongolia for the 20 years from 2001 to 2020, we used several datasets, including: MODIS 250 m reflectance image (MOD13Q1), MODIS 500 m land cover product (MCD12Q1), JRC Global Surface Water Mapping Layers Copernicus Global Land Service: Land Cover 100 m (CGL-S-LC100), Global PALSAR-2/PALSAR Forest/Non-Forest at 25 m, and the Shuttle Radar Topography Mission (SRTM). These data will be described in the following.

2.1.1. Satellite data for land cover classification

MOD13Q1 is the most popular input of remotely sensed images for automatically producing time series of land cover maps at medium resolution for many regions around the world (Clark et al. 2010; Nitze et al. 2015). We selected the data-source based on the following considerations: first, a sufficiently long time series must be available which is met either by Landsat or by MODIS data and second, the temporal resolution must be high enough to allow the creation of cloud-free mosaics over vast areas as Mongolia. Especially, the latter turned-out to be crucial in a previous study, which showed that high cloud coverage in Mongolia obstruct the creation of cloud-free mosaics using Landsat data (Phan et al. 2020). Furthermore, spatial resolution of the land cover product from which samples for classification are automatically selected should be in the same range as the multispectral data to reduce scale effects. Since samples have been selected from MODIS land cover products (described below), MODIS images at 250 m spatial resolution were used as predictors, because they feature a similar spatial resolution compared to the reference data. Therefore, in this study, we used MOD13Q1 as the main data source.

MOD13Q1 Version 6 is available since February 2000, with spatio-temporal resolutions of 16 days and 250 m. In addition, this data can be easily and freely accessed on the GEE platform. Phan et al. (2020) found for Mongolia, with a strong seasonal phenology, that a median composition of images from June to September produces classification results as good as the time series composition, but with much less computing cost. Therefore, in this study, we used two vegetation indices (NDVI & EVI) together with four spectral bands (red, near-infrared, blue, and mid-infrared) from June to September over 20 years (from 2001 to 2020) to composite 20 yearly median input images.

2.1.2. Land cover products for extracting samples

We used MCD12Q1 (500 m spatial resolution, available for each year from 2001 to 2019) to automatically generate reference samples for the land cover classification. The “Copernicus Global Land Service: Land Cover 100 m” (Buchhorn et al. 2021), which is available for years 2015 to 2019, has been used as an additional source.

In this study, the land cover types (i.e., land cover classes) to be discriminated in the new product were identified based on their spatial importance calculated from MODIS MDC12Q1. As shown in Figure A4, the dominant land cover types in Mongolia are grassland and bare land. These are also the land cover types we pay most attention in this study, since they play an important role in grassland monitoring and land degradation assessment over Mongolia. We reclassified land cover types into six classes, including water, forest, shrub land, grassland, bare land, and other land cover types such as wetland and built-up land.

Surface water bodies have been masked out using the “JRC Global Surface Water Mapping Layers” (Pekel et al. 2016). The reason to exclude water from the classification was the high spectral variability in Mongolia between and within seasons.

In addition, as shown in Figure A1, the forested area was underestimated in MODIS land cover products in comparison to the GlobeLand30 product. Therefore, we used the Global PALSAR-2/PALSAR Forest/Non-Forest at 25 m spatial resolution to select forest samples (Shimada et al. 2014).

2.1.3. Topographic data

In the literature, there are several studies that reported the improvement of classification results when integrating an auxiliary variable, such as those derived from the topography (Li et al. 2016; Jin et al. 2018, Phan et al. 2020). Therefore, in this study, we generated elevation, slope, aspect, east-ness and north-ness from the Shuttle Radar Topography Mission (SRTM) data and integrated them into the input data for classification models.

2.2. Methods

To produce yearly land cover maps based on a quality controlled but automatically derived training data set, 10 steps have been performed (Figure 1). The first 8 steps are necessary to derive the training data set. During step 9, the classification was performed. Since the Random Forest classifier was reported to be one of the best classifiers for land cover classification using remote sensing data in several studies (Thanh Noi and Kappas

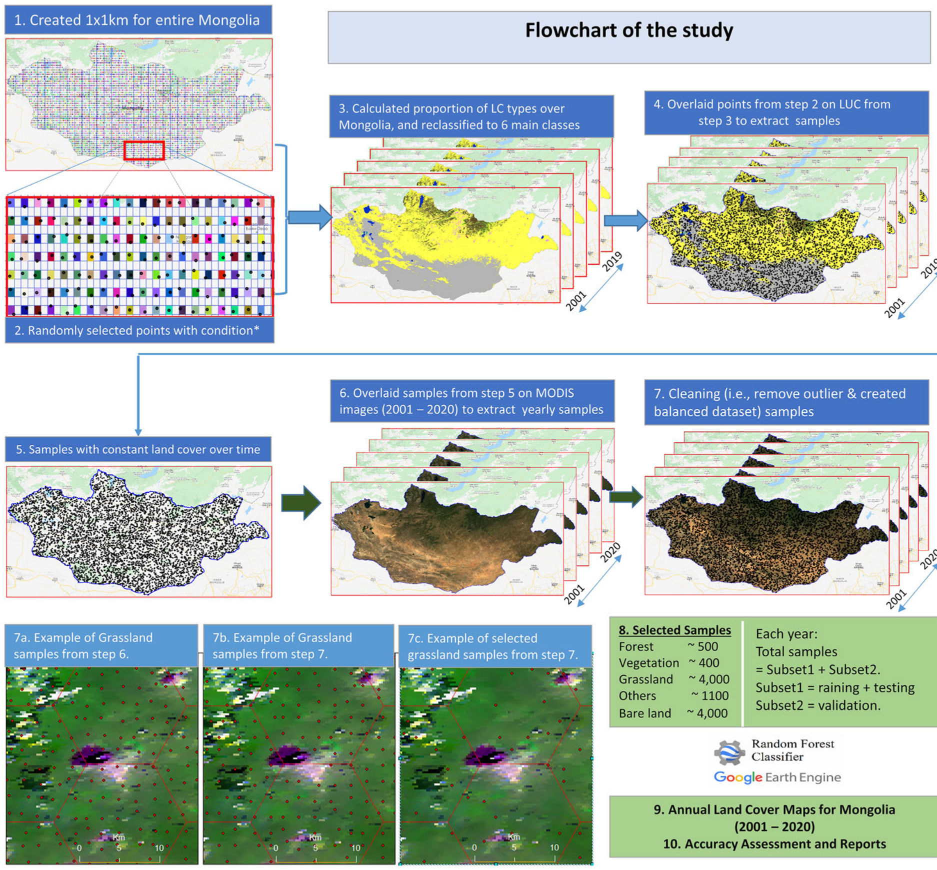

Figure 1. The flowchart of the study.

2017; Kelley et al. 2018; Teluguntla et al. 2018), it has been used as classifier for our classification. In the following, the 10 steps are described:

1. A 1x1 km grid system was created for entire Mongolia to make sure that the samples were well distributed throughout the country.
2. More than 140, 000 points within the grid cells of the grid system were randomly selected with the condition that if a grid cell is selected, the 8 neighbour grids must not be selected. This step was implemented to minimize spatial autocorrelation by ensuring that the samples were separated by a minimum distance of 1 km.
3. Based on the MODIS MCD12Q1 and the CGLS-LC100 land cover products, the proportion of each land cover type was calculated and the MCD12Q1 land cover product was reclassified into six land cover types. In our product, the five main groups of land cover types are: forest, shrub land, grassland, bare land and other land cover types, e.g., wetland and built-up land.
4. All points randomly selected in step 2 were overlaid above the LUC raster data from step 3 to extract a set of potential samples.
5. To minimize the error introduced by the usage of the land cover products, only potential samples have been further used, which were covered by a constant land

cover type throughout the investigation time from 2001 to 2020. Training samples for forest class have been selected based on the PALSAR-2/PALSAR Forest/Non-Forest product.

6. Annual mosaics for entire Mongolia were created from all available cloud-free MOD13Q1 images from June to September. If multiple images could contribute at a certain location, a median filter selected the respective reflectance values (Phan et al. 2020). From reflectance values of each yearly mosaic, two vegetation indices (NDVI, EVI) have been calculated. Finally, the reflectance values of the four bands from the MODIS mosaic, the two vegetation indices and topographic variables have been stacked. Using this stack, the training dataset has been derived for each location of the training samples.
7. Due to the low thematic accuracy of the reference data (i.e., accuracy of MCD12Q1 ranges between 60% and 90%, depending on regions), the samples extracted from the previous step contained outliers (Radoux et al. 2014). In this step, only those samples have been kept, whose reflectance values were located in the interquartile rangesof the respective class and satellite band. By this, we ensured a reduction in the number of samples not representative for the respective land cover class as well as reduction of imbalanced samples among land cover types.

In a second step, we further removed samples of the two most dominant classes, bare land and grasslands to guarantee a well-balanced dataset. Therefore, hexagons have been created covering the entire study area. Each hexagon was created so that it covers an area of 100 km^2. Within each hexagon, one sample at maximum has been selected for further processing.

Step 7 resulted in approximately 10,000 points for each year from 2001 and 2020. These samples were divided into two sub datasets with the proportional size of 60/40 for sub dataset 1 and 2, respectively. The most important requirement of this step was to keep proportions of land cover types in both samples equal. Sub dataset 2 was used for validation and sub dataset 1 was further divided into training and testing datasets with a proportion of 70% and 30%.

We used the Random Forest Classifier in Google Earth Engine with ntree = 500 to classify the land cover maps for 20 years, from 2001 to 2020. For each year, a separate model has been trained.

Each classified map was evaluated based on the testing, and the validation data directly in the GEE platform. Since Foody (2020) highly recommend that the kappa coefficient should not be used for land cover classification accuracy assessment, we used the overall accuracy (OA) and confusion matrix to evaluate our classification results.

Following the recommendations of Olofsson et al. (2014), independent samples should be used to further validate the land cover classification. Since we selected separate points with a minimum distance of 1 km for training, testing and validation (step8/Figure 1), we consider our validation datasets to be independent from the training dataset. However, both datasets inhibit similar errors depending on the accuracy of the MODIS land cover product. Consequently, a second dataset has been used for an external assessment. Therefore, validation samples from the GlobeLand30 product (30 m resolution) have been randomly collected to validate our land cover map in 2010.

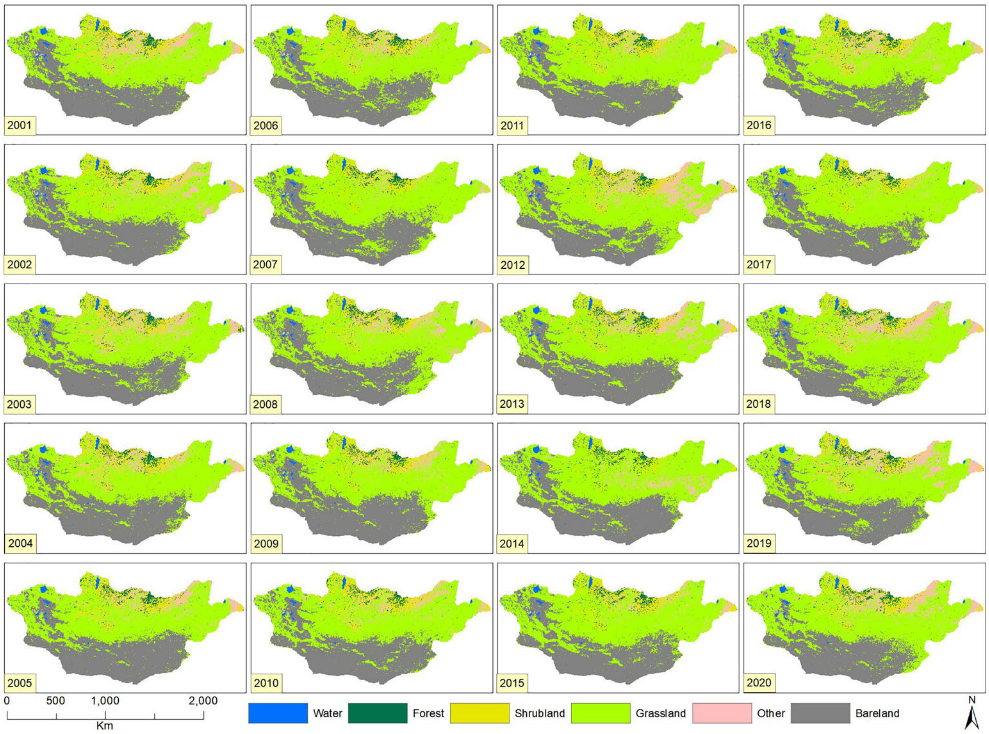

Figure 2. Spatial distribution of land cover types over Mongolia from the classification results.

3. Results

3.1. Land cover classification

Using the proposed method, we produced yearly land cover maps for entire Mongolia from 2001 to 2020 (Figure 2). The dominant land cover type over the last 20 years was grassland (on average, it accounts for 51.1%), followed by bare land (32.6%). The remaining classes (i.e., forest, and other) accounted for 16.3% of Mongolia.

As shown in Figure 2, the coverage of all land cover types varies both spatially and temporally. The most significant changes are related to grassland and bare land. The lowest and highest coverage of grassland was observed in 2001 and 2007, respectively. The lowest and highest coverage of bare land was observed in 2018 and 2009, respectively. This change mainly occurred in central to southeast Mongolia.

Due to the low proportion of some land cover types (including built-up land, wetland, and agriculture), their changes are hardly visible in the maps in Figure 2. Consequently, Figure 3 shows land cover transformations between years 2001, 2005, 2010, 2015, and 2020 in a graphical way, confirming that the main land cover changes occurred between grassland and bare land. In addition, it can be clearly seen that there were also changes between grassland and other class land cover types. This suggests that wetland, agricultural land, and grassland changed between years.

3.2. Classification accuracy

To validate the land cover map in 2010, samples from the GlobeLand30 product have been used as reference. The classification was in good agreement to this independent dataset with an overall

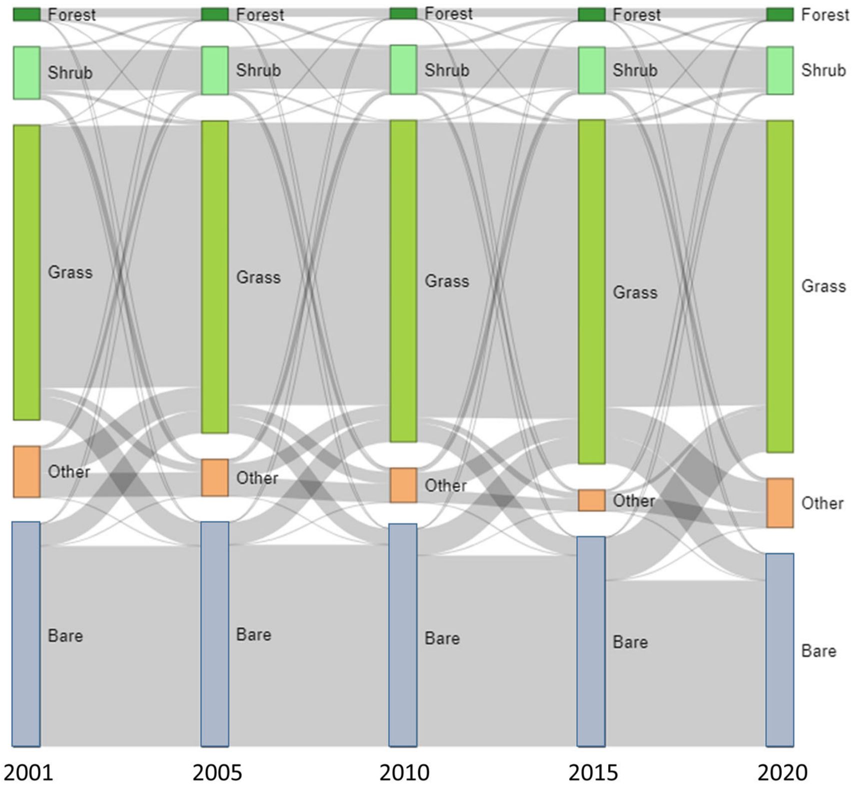

Figure 3. Land cover change in Mongolia based on selected years.

accuracy of 91.3% (Table 1). Highest errors have been observed between classes Grassland and Other, while only few classification errors in classes Bareland and Forest occurred.

For all other years between 2001 and 2020, validation rely on samples from the MODIS land cover product, which have been split into training, testing, and validation datasets. Our method minimized the spatial autocorrelation (to avoid overfitting), as well as ensured that the proportions of land cover type samples in the training, testing, and validation dataset are equal. The overall accuracy derived from these datasets was consistent over time (Table 2) and in the same order as the validation results against the fully independent dataset (Table 1).

3.3. The improvement of the land cover products

By comparing our land cover classification with the MODIS land cover product (MCD12Q1, reclassified), a large difference between these two products can be observed. In Figure 4, the differences between these land cover maps for 6 years are shown (2001, 2005, 2009, 2011, 2015, and 2019).

Since the land surface (i.e., land cover) in Mongolia varies during the year as well as from year to year, we visually compared the land cover maps with satellite images. To interpret and compare, as well as to identify the most accurate land cover maps (i.e., our classification versus the MODIS product), we selected 4 comparison locations (A, B, C, D: Figure 4). The false colour composition of MODIS reflectance bands (MOD13Q1, 250 m spatial resolution) in July of the corresponding year and the median composition of the

Table 1. Accuracy of the 2010 land cover classification based on GlobeLand30 product.

LC	Forest	Shrub	Grassland	Other	Bare land	Total	UA(%)
Forest	104	3	0	1	0	108	96.3
Shrub	2	58	0	0	0	60	96.7
Grassland	0	1	958	135	17	1111	86.2
Other	0	0	14	85	0	99	85.9
Bareland	0	0	7	0	689	696	99.0
Total	106	62	979	221	706	2074	1894.0
PA(%)	98.1	93.5	97.9	38.5	97.6	**OA = 91.3**	

Table 2. Overall accuracy of the classification.

Year	Training	Testing	Validation
2001	93.65	91.85	90.14
2002	91.78	90.80	90.76
2003	92.94	90.68	90.57
2004	93.88	92.75	91.95
2005	92.69	91.35	91.52
2006	92.90	91.11	91.68
2007	92.81	89.95	89.68
2008	92.20	89.24	90.11
2009	92.12	90.34	90.30
2010	92.18	89.39	89.76
2011	92.02	88.97	88.40
2012	92.34	91.41	90.66
2013	91.57	91.77	90.82
2014	91.06	87.04	87.23
2015	91.82	88.79	88.57
2016	91.02	87.97	87.71
2017	91.18	87.12	88.75
2018	92.33	89.81	90.03
2019	91.84	89.64	88.97
2020	92.36	90.73	90.59

Landsat images (image were selected between June and September of the corresponding years) were used for interpretation and comparison.

As shown in Figure 5, the MODIS products overestimate grassland area in all years. At the spatial zoom A and D, bare land was classified as grassland, at the zoom locations B and C, other land cover types such as forest, shrub land, and other classes were classified as grassland in the MODIS products. In other words, our classification land cover maps have much higher accuracy for grassland and the bare land classes. This information is important for Mongolia, because grassland and bare land information are of the utmost importance for monitoring greening and browning trends, relevant for grassland, and desertification monitoring over time.

4. Discussion

4.1. The role and methods to select quantity & quality of samples

The two big challenges for land cover mapping using remotely sensed data for large areas are: (i) large input image datasets need to be processed, and (ii) collecting samples of sufficient quantity and quality for such large areas is time-consuming (Loveland et al. 2000, Bartholomé and Belward 2005, Friedl et al. 2010, Gong et al. 2013, Zhang and Roy 2017).

Given the availability of the Google Earth Engine cloud-computing platform, the first issue is solved effectively, which has been reported in many studies (Gorelick et al. 2017; Phan et al. 2020; Tamiminia et al. 2020). Regarding the second issue, it is well-known that

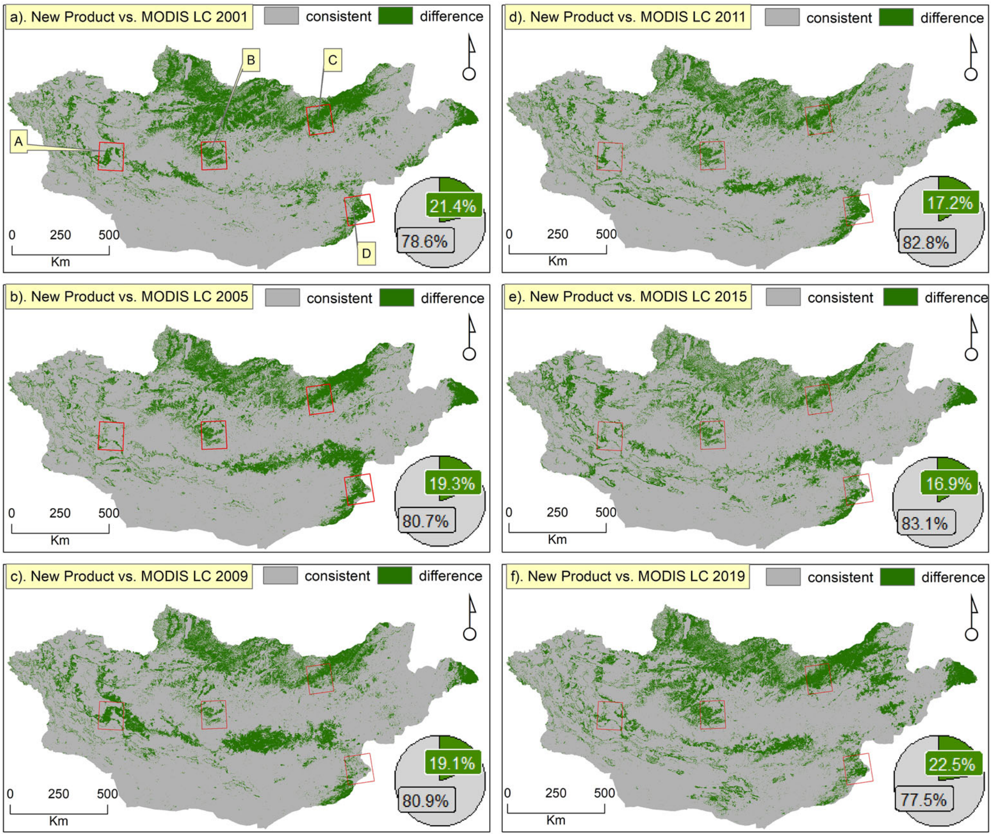

Figure 4. Difference between the new product and MODIS Land Cover Products in years (a). 2001; (b). 2005; (c). 2009; (d). 2011; (e). 2015; and (f). 2019. The pie charts show the percentage of consistent and different pixels between the classification and MODIS land cover products in the corresponding years.

training samples of sufficient quantity and high quality are a prerequisite to achieve classification results with high accuracy (Foody and Arora 1997; Persello and Bruzzone 2014; Li et al. 2021). For generation of samples, automatic methods have been developed (Xie et al. 2019). However, most of these studies only focused on the quantity (i.e., extracted high number of samples) while only few studies focused on the quality of samples. For instance, Xie et al. (2019) set four rules to select samples from MCD12Q1 land cover, but there were no rules to consider spatial autocorrelation. As a result, when evaluating the extracted samples, (i.e., training the model), they got a very high overall accuracy (i.e., 99.2%), but a low overall accuracy of 78% – 80% when these datasets were compared to independent datasets. Obviously, spatial autocorrelation in the samples led to overfitting.

In our results, the effects of the spatial autocorrelation were minimized by selecting point samples with a minimum distance of 1 km. The performance of this technique was confirmed in our results, i.e., the OA of training, testing, and validation datasets of all the classifications ranged from 89% to 91% (Table 2). However, not only the distance between sample points helped to achieve high and consistent OA values between sub-samples, but also that our method ensures to incorporate training samples all over the study region (i.e., that samples are well distributed over the study area). In addition, we minimized the effects of imbalanced samples between classes. The high OA was also constant through all the classifications of the 20 year results.

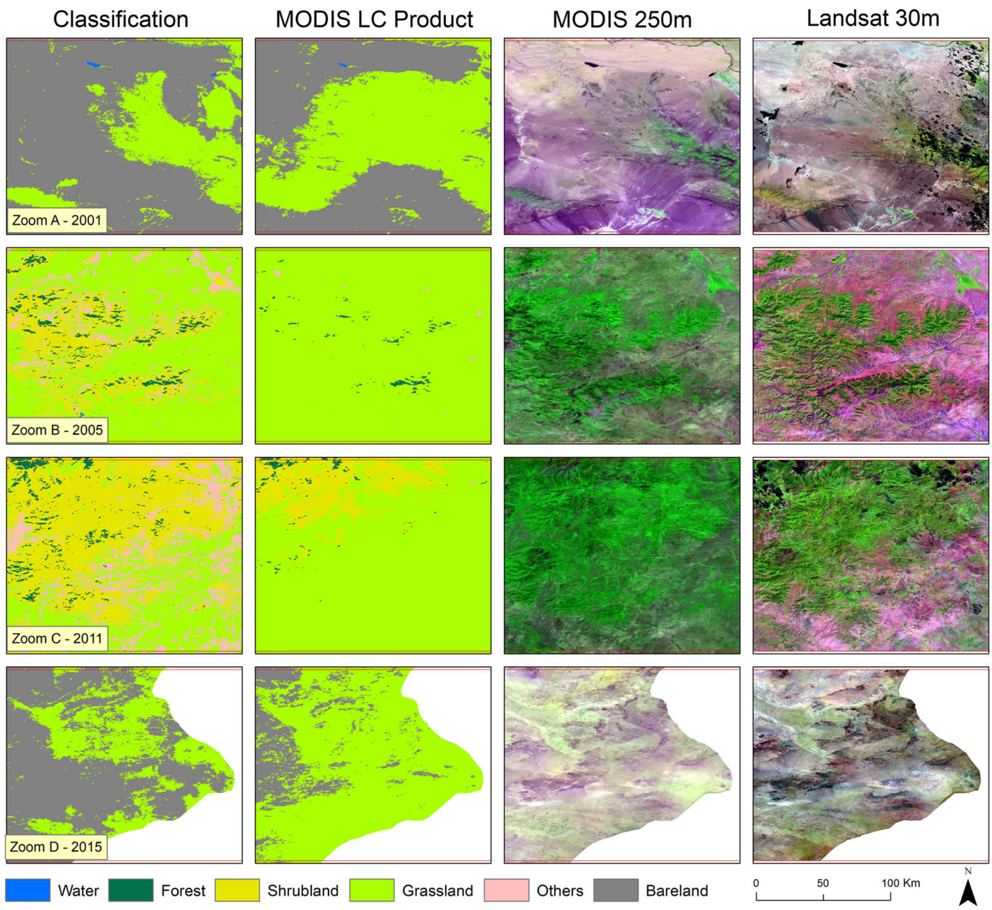

Figure 5. Spatial zooms (A–D) indicated in Figure 4 of the different products. The first column is the new product, the second column is MODIS LC product, third and fourth columns are the corresponding images from MODIS MOD13Q1 (250 m) and Landsat (30 m), respectively.

It is worth noting that, to our best knowledge, there is no study using an automatic method of collecting samples and considering that samples from all over the study region needs to be included in the samples. This issue has been raised in a recent publication (Li et al. 2021).

Another point that should be mentioned is that our analysis and calculations were mainly performed on the GEE platform. Therefore, the results demonstrated and confirmed that time series land cover maps (20 years) for the entire Mongolia (area ~ 1.564 million km^2) can be effectively analysed on the GEE platform.

4.2. Accuracy of the classifications

There are large differences between MODIS LC products and our classification results, and we could show that our classification results had higher accuracies and were more realistic than the MODIS LC product over Mongolia during the investigated time. This is consistent with the confusion matrix reports. For example, MODIS LC has an OA of approximately 75%, and our LC has an OA of approximately 90%. However, 75% from MODIS LC was calculated for global scale, meaning that there are areas that have a higher percentage than 75%, and vice versa. The OA of our classification was calculated for Mongolia only.

It is difficult to quantify the effects of the differences between MODIS LC and our product due to lacking adequate reference data. In fact, to our best knowledge, there is no study evaluating the performance of global land cover products over Mongolia. There is also no study quantifying how errors in the land cover data affect downstream applications in Mongolia. To quantify these effects, more studies are critically required, because they are key for the development of better strategies fighting climate change, food insecurity, or nature degradation.

The high accuracy of the classifications confirms that the proposed workflow is promising. Particularly, the thematic accuracy of the classified maps were much higher (i.e., 90% vs. 75%) than the original land cover data used as reference. This result is consistent with the study of Radoux et al. (2014).

4.3. Recommendation to automatically collect samples from existing land cover products

In the literature, particularly in the last decade, a number of studies have automatically collected samples for land cover classification using existing land cover products. There are also several studies that used the available land cover maps (usually of a higher spatial resolution than those of the corresponding study's classification results), to manually select samples for classification. However, in this study, we only discuss the automatic sample selection methods.

In general, most studies with automatically generated samples used coarser global land cover products to generate samples for finer resolution of classification results (Zhang and Roy 2017; Xie et al. 2019). Among these, MODIS land cover products were the most popular data sources for collecting samples for land cover classification at the higher resolutions such as Landsat, or even at 10 m spatial resolution (Zhang and Roy 2017, Xie et al. 2019).

However, three big issues that occur in automatically collected samples are rarely studied and discussed. These issues are spatial autocorrelation, imbalanced data, and distribution of the samples.

After automatically selecting samples, the samples have to be cleaned. In the literature, there are two popular methods for cleaning such samples. One method is based on the spatial resolution of the land cover reference map and the targeted land cover classification, in order to select consistent pixels (i.e., point samples). Another method is to remove pixels based on the outliers of spectral signatures (Radoux et al. 2014). The latter method is more widely applied than the first method, because it is difficult to collect enough samples for heterogeneous areas (such as urban areas). In our study, we also used the second method to clean the samples, because we focused more on the accuracy of grassland and bare land, which are quite specific in their spectral range of distribution.

Figure 5 shows the reason why cleaning samples is important and why we used the second cleaning approach in this study. If the input images for the reference land cover product were not carefully considered, the reference land cover product might not be able to distinguish between bare land and grassland. As a result, the yearly land cover products might be inaccurate, at least between grassland and bare land, as in our case. For example, if too many low red band (or NDVI) points are "fed" as input but labelled as grassland class, or in contrast, high NDVI points are labelled as bare land class, the trained model will be biased and feature lower accuracy in its prediction. That is why the samples should be consistent with the input reflectance images (by cleaning samples so that outliers are removed).

Beside cleaning the samples, we suggest that spatial resolutions of the reference land cover product and the new classification output should not be too much different(e.g., using 500 m land cover product to generate samples for 30 m classified land cover target). Even though this issue has not been widely studied, Roy and Kumar (2017) reported that "only about 5% of 1 km MODIS pixels over the Brazilian Tropical Moist Forest Biome (4 million km^2) contained homogeneous land cover mapped at 30 m". Therefore, it is not easy to find a 500 m land cover pixel that has single land cover types at 30 m resolution (Zhang and Roy 2017).

Thus, for applying automatic sample selecting methods, choosing approachable and feasible reference LC products for the targeted LC classification is an important and crucial step.

5. Conclusion

Land cover information has been defined as one of the most important pieces of information to understand the interaction between human activities, land surface, and atmosphere. Since remote sensing data got widely available, land cover maps are mainly produced based on satellite data, particularly if land cover maps are generated for regional to global scales. Due to the importance of land cover data, several land cover products have been developed at a global scale, and these products have been widely used for several downstream applications. However, the accuracy of the global land cover products in specific regions or at the national scales must receive more attention. A large number of studies have reported that the global land cover products do not perform equally well (i.e., different accuracy) from region to region. Furthermore, a number of recent studies also pointed out that the inaccuracy of the land cover products at the regional and/or national scale has a huge impact on downstream applications. Therefore, the accuracy of land cover products at the regional scale is of utmost importance, particularly regarding open access land cover products, which can be easily used in downstream applications. Land cover maps at national and/or regional scales including time series can now be produced for large and remote areas like Mongolia. The only possible way to generate samples is using the existing land cover products. Some common issues with automatically generated training samples are spatial autocorrelation, imbalanced land cover types, and insufficient distribution of samples. Our study is among the first studies using the automatic sample generating methods that consider all the mentioned issues. As a result, we achieved a very high accuracy (89-91%) over 20 years of land cover classifications. The accuracy of the method has been demonstrated not only in the OA reports, but also in the consistency between training, testing, and validating OA for all years. The accuracy was further proven with the additional assessment for the years 2010. We believe that our method will significantly contribute to future land cover classifications due to its simplicity, fastness, and accurateness. This method should be transferred and applied to other areas to further test it. If successful, it provides valuable data for many downstream applications in Mongolia and elsewhere.

Acknowledgments

This study was conducted within the framework of the MoreStep-Project ("Mobility at risk: Sustaining the Mongolian Steppe Ecosystem") and was funded by the German Federal Ministry of Education and Research (01LC1820B).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was funded by the German Federal Ministry of Education and Research (01LC1820B).

ORCID

Thanh-Noi Phan http://orcid.org/0000-0002-2747-5028

Data availability statement

The land cover maps from 2001 to 2020 have been submitd to https://www.pangaea.de/ (open access).

References

Ahmad S, Jia H, Chen Z, Li Q, Yin D, Israr M, Hayat W, H. Bilal, Ahmed R, Ashraf A. 2021. Effects of climate and land use changes on stream flow of Chitral river basin of northern highland Hindu-Kush region of Pakistan. J Hydro-Env Res. 38:53–62. doi:10.1016/j.jher.2021.08.001

Bai Y, Feng M, Jiang H, Wang J, Liu Y. 2015. Validation of land cover maps in china using a sampling-based labeling approach. Remote Sensing. 7(8):10589–10606.

Ban Y, Gong P, Giri C. 2015. Global land cover mapping using earth observation satellite data: recent progresses and challenges. ISPRS J Photogramm Remote Sens. 103:1–6.

Bartholomé E, Belward AS. 2005. Glc2000: a new approach to global land cover mapping from earth observation data. Int J Remote Sens. 26(9):1959–1977.

Buchhorn M, Bertels L, Smets B, De Roo B, Lesiv M, Tsendbazar N-E, Masiliunas D, Li L. 2021. Copernicus global land service: land cover 100m: version 3 Globe 2015-2019: algorithm theoretical basis document (Dataset v3.0, doc issue 3.4). Zenodo. doi:10.5281/zenodo.4723924

Clark ML, Aide TM, Grau HR, Riner G. 2010. A scalable approach to mapping annual land cover at 250 m using modis time series data: a case study in the dry chaco ecoregion of south america. Remote Sens Environ. 114(11):2816–2832.

Elmes A, Alemohammad H, Avery R, Caylor K, Eastman J, Fishgold L, Friedl M, Jain M, Kohli D, Laso Bayas J, et al. 2020. Accounting for training data error in machine learning applied to earth observations. Remote Sensing. 12(6):1034.

Foody GM. 2020. Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. Remote Sensing Environ. 239: 111630.

Foody GM, Arora MK. 1997. An evaluation of some factors affecting the accuracy of classification by an artificial neural network. Int J Remote Sens. 18(4):799–810.

Friedl MA, Sulla-Menashe D, Tan B, Schneider A, Ramankutty N, Sibley A, Huang X. 2010. Modis collection 5 global land cover: algorithm refinements and characterization of new datasets. Remote Sens Environ. 114(1):168–182.

Friedlingstein P, O'sullivan M, Jones MW, Andrew RM, Hauck J, Olsen A, Peters GP, Peters W, Pongratz J, Sitch S, et al. 2020. Global carbon budget 2020. Earth Syst Sci Data. 12(4):3269–3340.

Giri C, Zhu Z, Reed B. 2005. A comparative analysis of the global land cover 2000 and modis land cover data sets. Remote Sens Environ. 94(1):123–132.

Gong P, Wang J, Yu L, Zhao Y, Zhao Y, Liang L, Niu Z, Huang X, Fu H, Liu S, et al. 2013. Finer resolution observation and monitoring of global land cover: first mapping results with landsat tm and etm + data. Int J Remote Sens. 34(7):2607–2654.

Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. 2017. Google earth engine: planetary-scale geospatial analysis for everyone. Remote Sens Environ. 202:18–27.

Grekousis G, Mountrakis G, Kavouras M. 2015. An overview of 21 global and 43 regional land-cover mapping products. Int J Remote Sens. 36(21):5309–5335.

Hermosilla T, Wulder MA, White JC, Coops NC. 2022. Land cover classification in an era of big and open data: optimizing localized implementation and training data selection to improve mapping outcomes. Remote Sens Environ. 268:112780.

Jin Y, Liu X, Chen Y, Liang X. 2018. Land-cover mapping using random forest classification and incorporating ndvi time-series and texture: a case study of central Shandong. Intern J Remote Sensing. 39(23): 8703–8723.

Kelley LC, Pitcher L, Bacon C. 2018. Using google earth engine to map complex shade-grown coffee landscapes in northern Nicaragua. Remote Sensing. 10(6):952.

Li X, Chen W, Cheng X, Wang L. 2016. A comparison of machine learning algorithms for mapping of complex surface-mined and agricultural landscapes using ziyuan-3 stereo satellite imagery. Remote Sensing. 8(6):514.

Li H, Ding J, Zhang J, Yang Z, Yang B, Zhu Q, Peng C. 2020. Effects of land cover changes on net primary productivity in the terrestrial ecosystems of china from 2001 to 2012. Land. 9(12):480.

Li C, Ma Z, Wang L, Yu W, Tan D, Gao B, Feng Q, Guo H, Zhao Y. 2021. Improving the accuracy of land cover mapping by distributing training samples. Remote Sensing. 13(22):4594.

Liu Q, Zhang Y, Liu L, Li L, Qi W. 2019. The spatial local accuracy of land cover datasets over the qiangtang plateau, high asia. J Geogr Sci. 29(11):1841–1858.

Loveland TR, Reed BC, Brown JF, Ohlen DO, Zhu Z, Yang L, Merchant JW. 2000. Development of a global land cover characteristics database and igbp discover from 1 km avhrr data. Int J Remote Sens. 21(6-7):1303–1330.

Majasalmi T, Rautiainen M. 2021. Representation of tree cover in global land cover products: Finland as a case study area. Environ Monit Assess. 193(3):121. Available from: https://www.ncbi.nlm.nih.gov/pubmed/33576910.

Nitze I, Barrett B, Cawkwell F. 2015. Temporal optimisation of image acquisition for land cover classification with random forest and modis time-series. Int J Appl Earth Obs Geoinf. 34:136–146.

Olofsson P, Foody GM, Herold M, Stehman SV, Woodcock CE, Wulder MA. 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sensing Environ. 148:42–57.

Pekel JF, Cottam A, Gorelick N, Belward AS. 2016. High-resolution mapping of global surface water and its long-term changes. Nature. 540(7633):418–422. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27926733.

Pérez-Hoyos A, Rembold F, Kerdiles H, Gallego J. 2017. Comparison of global land cover datasets for cropland monitoring. Remote Sensing. 9(11):1118.

Persello C, Bruzzone L. 2014. Active and semisupervised learning for the classification of remote sensing images. IEEE Trans Geosci Remote Sensing. 52(11):6937–6956.

Phan TN, Kuch V, Lehnert LW. 2020. Land cover classification using google earth engine and random forest classifier—the role of image composition. Remote Sensing. 12(15):2411.

Pongratz J, Schwingshackl C, Bultan S, Obermeier W, Havermann F, Guo S. 2021. Land use effects on climate: current state, recent progress, and emerging topics. Curr Clim Change Rep. 7(4):99–120.

Radoux J, Lamarche C, Van Bogaert E, Bontemps S, Brockmann C, Defourny P. 2014. Automated training sample extraction for global land cover mapping. Remote Sensing. 6(5):3965–3987.

Roy DP, Kumar SS. 2017. Multi-year modis active fire type classification over the Brazilian tropical moist forest biome. Int J Digital Earth. 10(1):54–84.

Shimada M, Itoh T, Motooka T, Watanabe M, Shiraishi T, Thapa R, Lucas R. 2014. New global forest/non-forest maps from alos palsar data (2007–2010). Remote Sens Environ. 155:13–31.

Siddiqui A, Kushwaha G, Nikam B, Srivastav SK, Shelar A, Kumar P. 2021. Analysing the day/night seasonal and annual changes and trends in land surface temperature and surface urban heat island intensity (suhii) for Indian cities. Sustainable Cities Society. 75:103374.

Tamiminia H, Salehi B, Mahdianpari M, Quackenbush L, Adeli S, Brisco B. 2020. Google earth engine for geo-big data applications: a meta-analysis and systematic review. ISPRS J Photogramm Remote Sens. 164:152–170.

Tateishi R, Hoan NT, Kobayashi T, Alsaaideh B, Tana G, Phong DX. 2014. Production of global land cover data—glcnmo2008. JGG. 6(3):99–122. doi:10.5539/jgg.v6n3p99.

Teluguntla P, Thenkabail PS, Oliphant A, Xiong J, Gumma MK, Congalton RG, Yadav K, Huete A. 2018. A 30-m landsat-derived cropland extent product of Australia and china using random forest machine learning algorithm on google earth engine cloud computing platform. ISPRS J Photogramm Remote Sens. 144:325–340.

Thanh Noi P, Kappas M. 2017. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery. Sensors (Basel). 18(2):18. Available from: https://www.ncbi.nlm.nih.gov/pubmed/29271909.

Tulbure MG, Hostert P, Kuemmerle T, Broich M, Disney M, Hernández-Stefanoni J. 2021. Regional matters: on the usefulness of regional land-cover datasets in times of global change. Remote Sens Ecol Conserv. doi:10.1002/rse2.248 (2021).

Wessels K, Van Den Bergh F, Roy D, Salmon B, Steenkamp K, Macalister B, Swanepoel D, Jewitt D. 2016. Rapid land cover map updates using change detection and robust random forest classifiers. Remote Sensing. 8(11):888.

Xie S, Liu L, Zhang X, Yang J, Chen X, Gao Y. 2019. Automatic land-cover mapping using landsat time-series data based on google earth engine. Remote Sensing. 11(24):3023.

Yang Y, Xiao P, Feng X, Li H. 2017. Accuracy assessment of seven global land cover datasets over china. ISPRS J Photogramm Remote Sens. 125:156–173.

Zhang X, Liu L, Chen X, Gao Y, Xie S, Mi J. 2021. Glc_fcs30: global land-cover product with fine classification system at 30 m using time-series landsat imagery. Earth Syst Sci Data. 13(6):2753–2776.

Zhang HK, Roy DP. 2017. Using the 500 m modis land cover product to derive a consistent continental scale 30 m landsat land cover classification. Remote Sens Environ. 197:15–34.

Zheng Q-H, Chen W, Li S-L, Yu L, Zhang X, Liu L-F, Singh RP, Liu C-Q. 2021. Accuracy comparison and driving factor analysis of lulc changes using multi-source time-series remote sensing data in a coastal area. Ecol Inf. 66:101457.

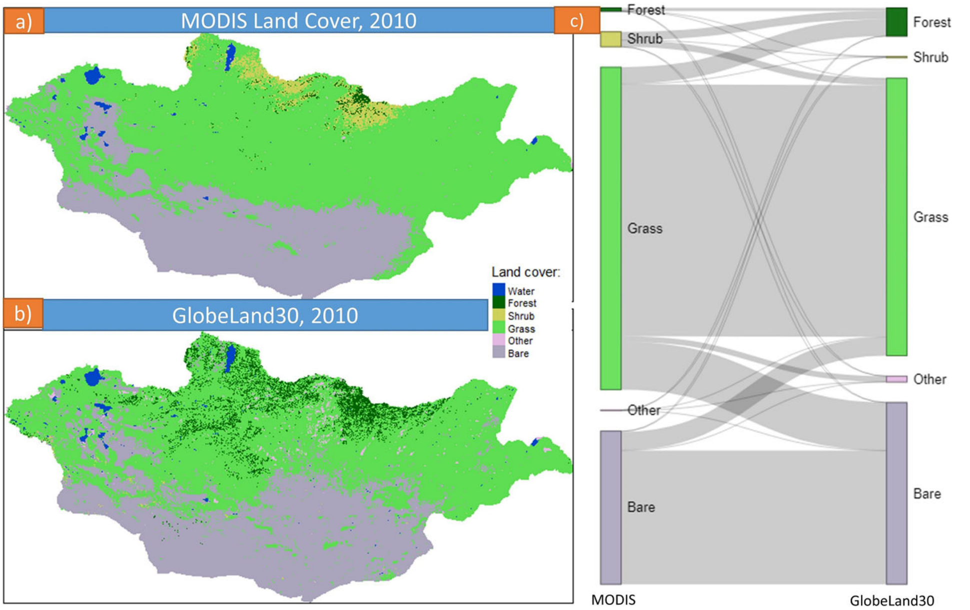

Figure A1. A simple comparison between two LC products (reclassified & resampled) over Mongolia: a) MODIS MCD12Q1 land cover, b) GlobeLand30, and c) the difference between these two land cover products.

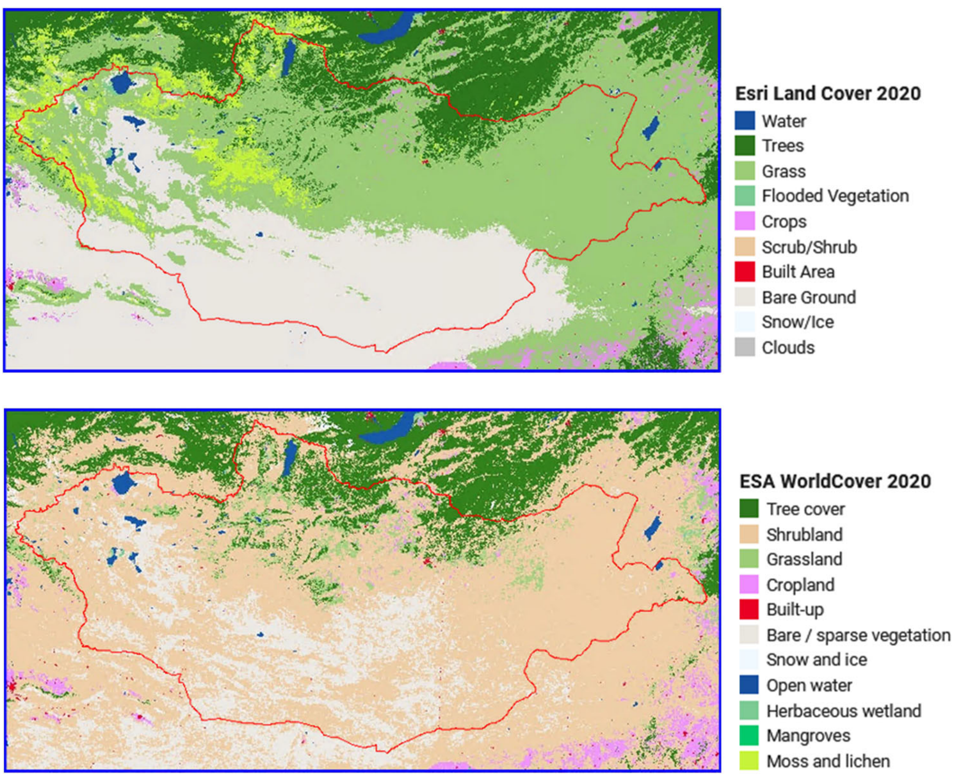

Figure A2. Shows the difference between two land cover products ESRI Land Cover 2020 versus ESA Land Cover 2020

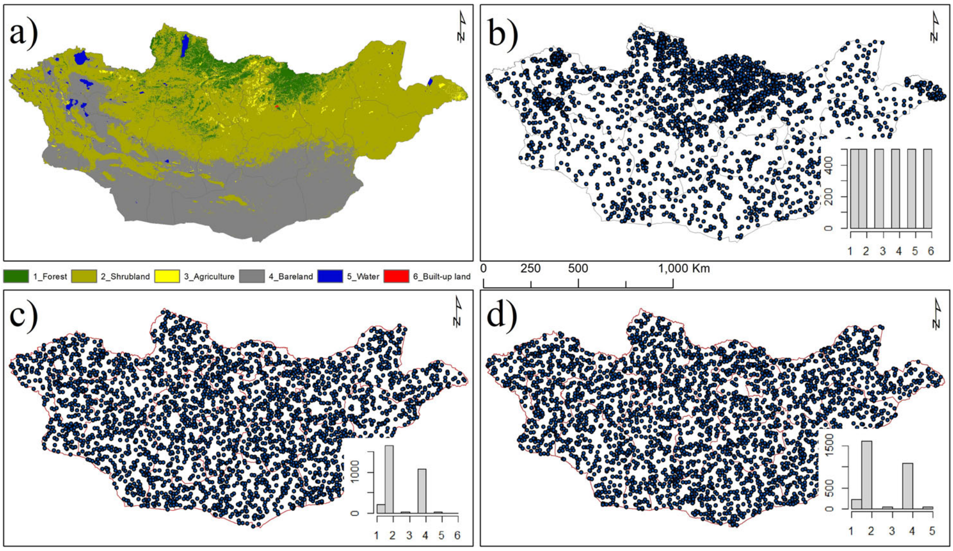

Figure A3. Differences between popular methods used for automatically collecting samples. a) Proba-V Land Cover distribution over Mongolia; b) Random and equal selection of 500 sample points for each of the 6 land cover types; c) Random selection of 3000 sample points; d) Same as c) but with minimum distance of 1 km between points. The histograms at the lower right corner show the number of samples for each land-use class.

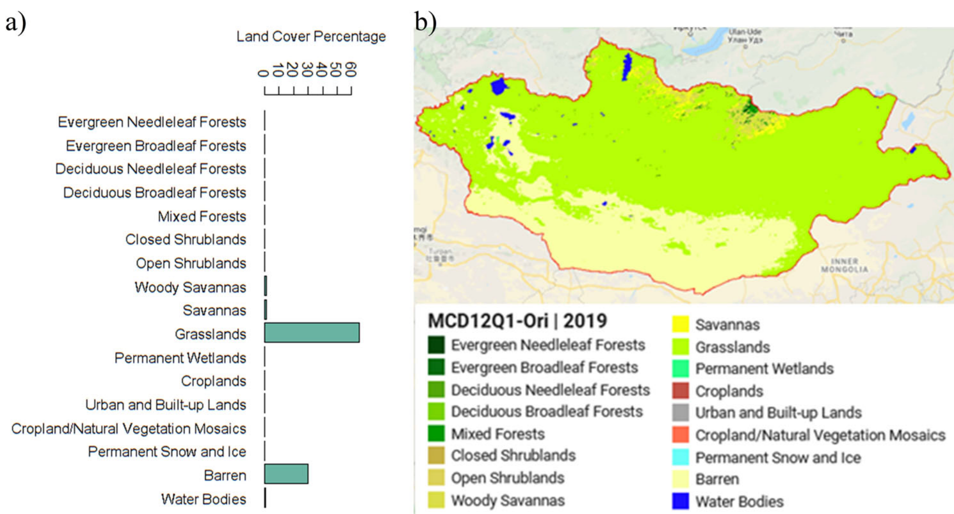

Figure A4. Land cover type in Mongolia based on MODIS LC product MCD12Q1 in 2019. a) Percentage of land cover types; b) spatial distribution of land cover type over Mongolia.