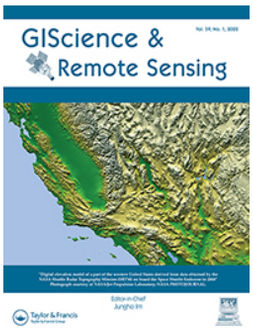

GIScience & Remote Sensing

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tgrs20

Improved forest cover mapping by harmonizing multiple land cover products over China

Shili Meng, Yong Pang, Chengquan Huang & Zengyuan Li

To cite this article: Shili Meng, Yong Pang, Chengquan Huang & Zengyuan Li (2022) Improved forest cover mapping by harmonizing multiple land cover products over China, GIScience & Remote Sensing, 59:1, 1570-1597, DOI: 10.1080/15481603.2022.2124044

To link to this article: https://doi.org/10.1080/15481603.2022.2124044

Published online: 27 Sep 2022.

GISCIENCE & REMOTE SENSING
2022, VOL. 59, NO. 1, 1570–1597
https://doi.org/10.1080/15481603.2022.2124044

Taylor & Francis
Taylor & Francis Group

Improved forest cover mapping by harmonizing multiple land cover products over China

Shili Meng[a,b], Yong Pang[a,b], Chengquan Huang[c] and Zengyuan Li[a,b]

[a]Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing, China; [b]Key Laboratory of Forestry Remote Sensing and Information System, National Forestry and Grassland Administration, Beijing, China; [c]Department of Geographical Sciences, University of Maryland College Park, MD, USA

ABSTRACT

Fine resolution land cover products are becoming increasingly more available for many regions. These products, however, may not meet the quality requirements of many applications. This study provides an approach for improving land cover mapping by leveraging existing products and clear view Landsat composites. Assessments using independent reference datasets revealed that the CAF-LC30 2020 product derived using this approach over China was more accurate than four existing land cover products. Its overall accuracy with field observations was 2.94% to 10.28% higher than those of the four existing land cover products in northeast China and was 2.10% to 8.18% better across China. It provided a more accurate representation of the land cover types in many regions where the existing land cover products had large classification errors. Forest areas calculated using the CAF-LC30 2020 for the 31 provinces and autonomous regions and municipalities (PARM) in mainland China were better correlated with those reported by the most recent National Forest Inventory (NFI) survey than areas calculated using the other four existing land cover products. Therefore, the CAF-LC30 2020 product should be a better alternative for understanding China's forests in 2020 than the other four existing land cover products.

ARTICLE HISTORY
Received 16 April 2022
Accepted 8 September 2022

KEYWORDS
Land cover product; forest; integration; landsat; China

Introduction

As a key descriptor of the Earth's surface, land cover (LC) is highly relevant for global change research and a wide range of applications (Townshend 1994). While many earth system processes are affected or driven by the spatial distribution and/or temporal dynamics of land cover, the latter is also shaped by both natural processes and human activities (Foley et al. 2005). Therefore, land cover products are among the basic datasets needed to understand drivers and processes of global change (Wulder et al. 2018), and to advance research and applications in ecology, environmental sciences, resource management, and sustainable development (Townshend et al. 2012; Andrew, Wulder, and Nelson 2014).

Since the launch of the first Landsat satellite in the 1970s, remote sensing has become a primary data source for land cover research (Townshend 1992; Foody 2002). Thanks to advances in remote sensing data source, algorithm research, and computing power, land cover mapping capabilities have evolved from local to national and global scales (Hansen and Loveland 2012). Following pioneering efforts to map global land cover at 1-degree (DeFries, Hansen, and Townshend 1995) and 8 km resolutions (DeFries et al. 1998), several land cover products were developed at 1 km or sub-km resolutions using observations acquired by the Advanced Very High Resolution Radiometer (AVHRR) (Hansen et al. 2000; Loveland et al. 2000), Moderate Resolution Imaging Spectroradiometer (MODIS) (Friedl et al. 2002, 2010; Tateishi et al. 2011), VEGETATION (Bartholome and Belward 2005), and Medium Resolution Imaging Spectrometer (MERIS) (Arino et al. 2007, 2008; Bicheron et al. 2008). While these coarse-resolution products are highly valuable to support climate and other models at the global scale, they cannot provide useful information for many individual land cover patches related to human activities (e.g. urban, crop fields, ponds, etc.), which are often smaller than the resolutions of these land cover products.

Efforts to create global Landsat datasets through NASA's Global Land Survey (GLS) program for public use (Tucker, Grant, and Dykstra 2004; Gutman et al.

CONTACT Yong Pang pangy@ifrit.ac.cn

2008, 2013) ushered opportunities to achieve sub-hectare spatial resolutions in global land cover mapping. Example products developed based on GLS datasets included early Landsat-based global land cover maps (Chen et al. 2015; Gong et al. 2013), forest cover change products (Feng et al. 2016; Kim et al. 2014; Sexton et al. 2015, 2013), and urban imperviousness products (Brown de Colstoun et al. 2017). The combination of free data policies adopted by the Landsat (Woodcock et al. 2008; Wulder et al. 2012; Zhu et al. 2019), Sentinel-1, and Sentinel-2 programs, and low cost or no-cost access to powerful cloud computing systems like the Google Earth Engine (GEE), Amazon Web Services (AWS), Microsoft Azure, and Descartes Labs (Yang et al. 2017) greatly improved the feasibility to generate land cover products at 10 to 30 m spatial resolutions on global scale. During the past decade, Hansen et al. (2013) used GEE and Landsat data to map forest extent and forest cover change since 2000. Pekel et al. (2016) quantify the changes of global surface water from 1984 to 2015 by the inventories and national descriptions, statistical extrapolation, and satellite imagery. Zhang et al. (2019, 2021) produced a suite of global 30 m fine classification maps (GLC_FCS30) using the globally feature spectral library. Liu et al. (2021) developed global 30 m annual-to-seasonal land cover maps for 1985–2020 with about 80% overall accuracy at first-level classification. By leveraging the finer spatial resolutions provided by Sentinel-1 and -2 than Landsat, Gong et al. (2019) produced a global land cover product at the 10 m resolution for the year 2017. Two other global 10 m land cover products were also released recently, one by the European Space Agency (ESA) (Zanaga et al. 2021) that mapped for the year 2020 using the Sentinel-1 and Sentinel-2 data, this product contained 11 land cover types and got an overall accuracy of 74.4% globally; the other by the Environmental Systems Research Institute (ESRI) (Karra et al. 2021) which was derived from Sentinel-2 images in 2020 by deep learning classification algorithm, this product contained 10 classes and achieved 85.9% overall accuracy on the validation set.

While the proliferation of global fine resolution land cover products is an indicator of good progress in land cover research, having multiple products over the same area can be a source of confusion for many users, because there are often substantial disagreements among those products. For example, Song, Huang, and Townshend (2017) analyzed 6 global land cover products with spatial resolutions ranging from 250 m to 1 km and found that disagreements among these products were high regarding the distribution of forests. Fritz, See, and Rembold (2010) showed that different land cover products not only differed substantially at the pixel level, they also resulted in very different area estimates for arable land at national to continental scales. These disagreements are indicators of uncertainties with at least some of the existing land cover products. Such uncertainties often resulted in excessive spurious changes when multi-temporal land cover products were compared to identify land cover change (Friedl et al. 2010; Friedl and Sulla-Menashe 2019). One approach to reduce those uncertainties is to produce a new product by using better training data, more accurate classification algorithms, less noisy inputs, and more predictor variables that can improve class separability. Alternatively, differences and similarities among existing land cover products may also be useful for improving land cover mapping. For example, Jung et al. (2006) created a joint 1 km land cover map by integrating multiple global land cover products. Fritz et al. (2011) developed an approach for harmonizing multiple land cover products, including Global Land Cover Map (GLC-2000) (Bartholomé and Belward, 2005), MODIS Land Cover product (MOD12V1) (Friedl et al. 2002), GlobCover (Bicheron et al. 2008), MODIS Crop Likelihood (Pittman et al. 2010), and AFRICover dataset (https://www.fao.org/3/bd854e/bd854e.pdf), to improve cropland mapping at the 1 km resolution for Africa. Song et al. (2014) developed a machine learning-based data fusion approach for harmonizing several land cover datasets to improve forest cover mapping at the 5 km spatial resolutions.

The primary goal of this study was to develop an approach that can leverage the differences and complementarity among existing land cover products to improve forest cover mapping at sub-hectare spatial resolutions. In this approach, multiple existing land cover products from 2000 to 2020 (i.e. GlobeLand30, GLC_FCS30, and ChinaCover, see Table 1 for more details) were harmonized and used to define an intersection map that included pixels where those products agreed with one another regarding the land cover types for those pixels. We demonstrated that the class labels of these pixels were highly accurate for several classes, and hence used a fraction of them

Table 1. List of 30 m land cover products used as input to the land cover mapping approach.

Product name	Year of available data layers				Input Data	Classification system	First-level land cover type	Organization	Source
GlobeLand30	2000	2010		2020	Landsat-TM /ETM +/OLI HJ-1 A/B	10 first-level classes	cultivated land, forest, grassland, shrubland, wetland, water bodies, tundra, artificial surfaces, bareland, permanent snow and ice	National Geomatics Center of China	http://www.globallandcover.com/home_en.html
GLC_FCS30	2000		2015	2020	Landsat-TM /ETM +/OLI	30 second-level classes	\	Chinese Academy of Science	https://data.casearth.cn/en/
ChinaCover	2000	2010	2015		Landsat-TM /ETM+ HJ-1 A/B	6 first-level 38 second-level classes	forest lands, grasslands, croplands, wetlands, built-up lands, other lands	Chinese Academy of Science	http://www.geodata.cn/data/index.html?ownername = %E5%90%B4%E7%82%B3%E6%96%B9

to train machine learning algorithms needed to classify clear view images created using an improved image compositing method. The existing land cover products were also used to define the maximum spatial extent of the forest, which was then used to reduce confusion between forest and other land cover types during post-processing. We used this approach to produce a China-wide 30 m forest cover map for 2020 (CAF-LC30 2020) and conducted comprehensive assessments of this product and four other land cover products developed for 2020.

Data

Landsat images

In this study, we used Landsat 8 OLI surface reflectance images available from the GEE platform that had cloud coverage <70%. The majority of the images were acquired in 2020. However, there were a few extremely cloudy regions where no cloud-free observations were available in 2020, and 2019 or/and 2021 images were included to create clear view composites. These regions were small in general, together they accounted for less than 5% of the total area of China.

Existing land cover products

Three sets of land cover products (i.e. GlobeLand30, GLC_FCS30, and ChinaCover) were used as input to the mapping algorithm. Each set consisted of land cover maps produced for three different years (Table 1). These land cover products were developed by different groups using different classification systems. (1) The GlobeLand30 dataset has 10 broad land cover types (Chen et al. 2015), and multiple assessment studies showed that overall accuracy of over 80% was achieved (Brovelli et al. 2015; Arsanjani, Tayyebi, and Vaz 2016; Chen and Chen 2018); (2) Developed based on a DataCube data structure and a global spectral library SPECLib (spatial-temporal spectral library), the GLC_FCS30 product suite had a reported accuracy of 82.5% (Zhang et al. 2019, 2021). This product suite separated some broad classes into more detailed classes in certain regions, resulting in a total of 30 classes; (3) The ChinaCover dataset used a classification scheme designed based on the classification system developed by United Nations Framework Convention on Climate Change (UNFCCC), which included 6 first-level classes and 38 second-level classes, the reported accuracies were 94% for first-level and 86% for second-level on independent random sampling evaluation (Wu et al. 2017).

As we started to evaluate the CAF-LC30 2020 product derived through this study near the end of our project, two new land cover products for 2020 became available. One was the China Land Cover Dataset (CLCD-2020) generated by Wuhan University (Yang et al. 2021). The other was the 10 m WorldCover land cover map produced by ESA (Zanaga et al. 2021), hereinafter referred to as ESA10-2020. We used the reference data collected through this study to provide an independent assessment of the quality of both products in China.

Reference data

Two sets of reference data were used to provide pixel-level assessments of the CAF-LC30 2020 map and four

other existing maps developed for 2020, including GlobeLand30-2020, GLC_FCS30-2020, CLCD-2020, and ESA10-2020.

The first set included field data collected in 2020 over 8,895 field survey plot locations distributed throughout the three provinces in northeastern China, including Heilongjiang, Jilin, and Liaoning. This dataset (referred to as the NE-China reference dataset hereafter) also covered a portion of eastern Inner Mongolia (Figure 1 (right)). The plots in this dataset were located within 2 km of accessible roads in order to achieve acceptable data collection efficiencies. However, several requirements were defined upfront to ensure that the selected plots were distributed across northeastern China as evenly as possible and that the collected data were representative. For example, no more than one plot should be selected from within one land cover object (e.g. a large reservoir or a large, homogeneous forest patch), nor should two plots having the same land cover type be located within 2 km of each other. To ensure that all counties were represented, every county should have at least 210 plots. To reduce potential ambiguities in determining the land cover type at any given location, plots were selected from homogeneous areas no smaller than 3,600 m^2, preferably larger than 10,000 m^2. Field data collected at each plot location included GPS coordinates, field photos, and descriptions of surface conditions and land cover type. Seven land cover types were identified in the field, including cultivated land, forest land, grassland, wetland, water, artificial surface, and bare land.

The second reference dataset included permanent plots selected by the Chinese Terrestrial Ecosystem Research Network (CTERN), which consisted of ecological stations established to represent different ecosystem types, including forest, cropland, grassland, shrubland, wetland, water, and urban (Figure 1 (left)). A CTERN plot has a square shape with an area ranging from 0.04 km^2 to 0.06 km^2. As of this study, the CTERN dataset consisted of mostly forest plots. Key factors considered in selecting CTERN plots included moisture and temperature gradients, forest type, and site quality. Further, all forested plots were collocated with the grids used by the National Forest Inventory (NFI) program to ensure that CTERN stations were distributed as evenly as possible. Nationally representative CTERN station networks for nonforest ecosystems have yet to be established.

The NE-China and CTERN datasets were highly complementary for evaluating the forest cover map developed through this study. The NE-China dataset provided a dense sampling of the entire northeastern China. The sample sizes for most classes were

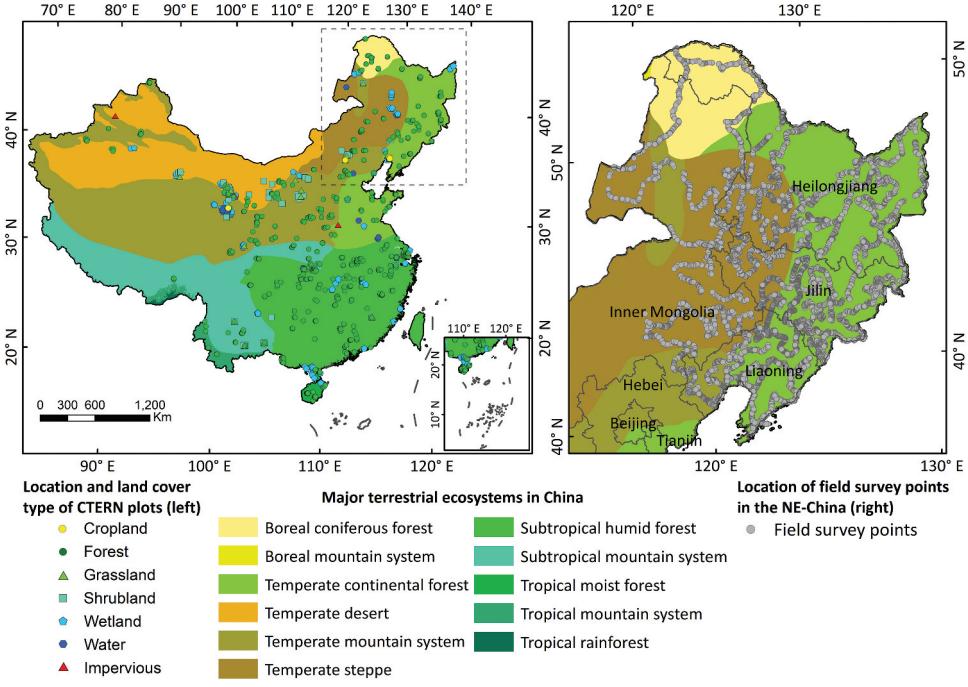

Figure 1. Spatial distribution of CTERN plots available from 2017 to 2020 (left) and plots in the NE-China reference dataset (right).

Table 2. Number of field plots to provide pixel-level assessments of the CAF-LC30 2020 and four existing land cover products developed for 2020.

Reference dataset name	Cropland	Forest	Grassland	Wetland	Water	Impervious	Bareland	Total
NE-China	4,120	2,377	474	349	341	1,158	76	8,895
CTERN	5	2,294	70	102	315	23	4	2,813
Total	4,125	4,671	544	451	656	1,181	80	11,708

adequate for deriving agreement metrics with satisfactory precision for northeastern China (Table 2). While the CTERN dataset had very few plots for cropland, bareland, and impervious surfaces, it included more than 2,000 plots representing most forest ecosystems in China, and therefore was ideal for evaluating how well those forests were mapped in this study.

In addition to the above reference datasets, we also obtained provincial-level forest area estimates reported by the National Forestry and Grassland Administration (NFGA) of China. These estimates were derived based on inventory data collected through China's NFI program. Established in the 1970s, the NFI program produces forest inventory data once every five years. The 9th NFI was conducted between 2014 and 2018, providing forest area estimates for 31 provinces and autonomous regions and municipalities (PARM) in mainland China (Table A2). These estimates were used to evaluate forest areas calculated from different land cover products, including the CAF-LC30 2020 map produced through this study and the four existing land cover products available for 2020. The 9th NFI revealed that in 2014–2018, the national forest coverage was 22.96%. In addition to both natural and planted forests, forest areas considered by the NFI program also included arbor forests and bamboo forests with canopy density above 20%. In certain regions (most of the semi-arid/arid areas in northwestern China), some shrub areas with canopy density above 30% were also classified as forest land.

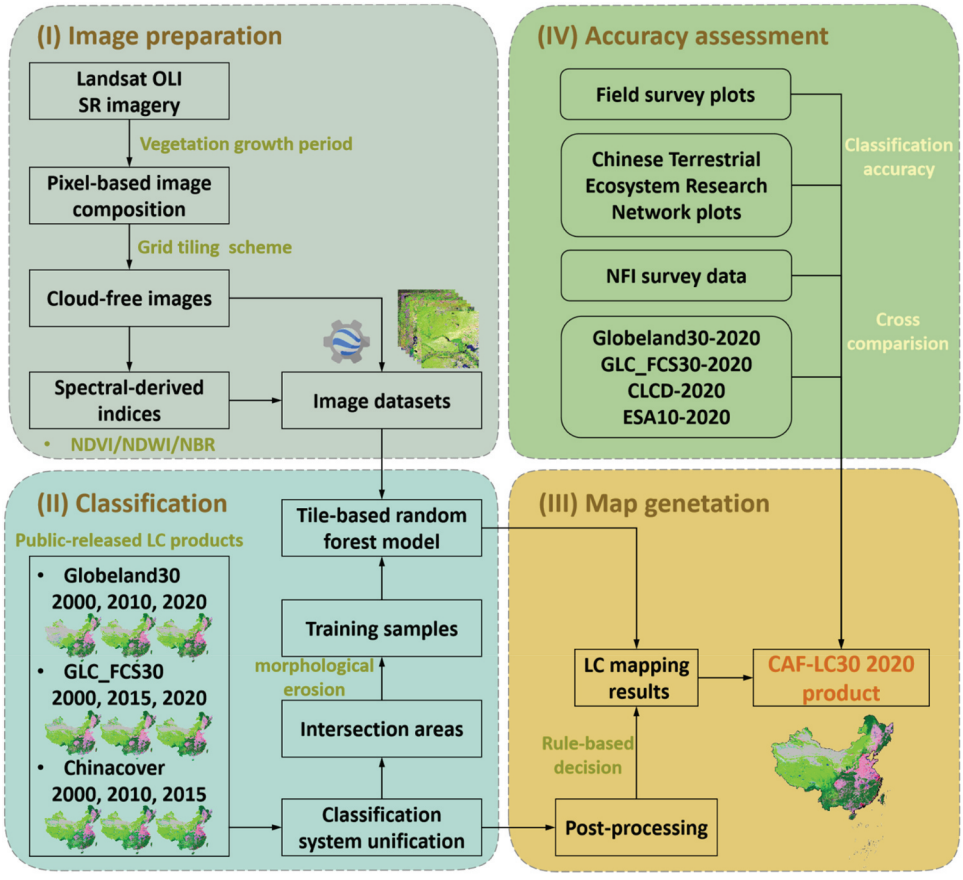

Figure 2. The workflow for developing and evaluating the CAF-LC30 2020 product.

Methods

The approach for producing the CAF-LC30 2020 map consisted of four major steps, including i) image preparation, ii) initial land cover classification; iii) post-processing and final map generation; and iv) accuracy assessment (Figure 2). The final product and four existing land cover products available for 2020 were evaluated using the reference datasets.

Images preparation

The primary goal of this step was to produce clear view image composites for 2020. Due to frequent cloud cover in most regions, obtaining cloud-free images for any given area is often challenging. Multi-temporal compositing is a common approach for generating cloud-free imagery products for large areas (Roy et al. 2010; Griffiths et al. 2013; White et al. 2014; Hermosilla et al. 2015; Zhu et al. 2015; Wulder et al. 2016). In this study, we followed the best available pixel (BAP) algorithm proposed by White et al. (2014) to generate cloud-free imagery. This algorithm uses a suite of scores to represent the quality of observations acquired in different date pixels, including the day-of-year in the leaf-on season, opacity, and the distance to cloud and cloud shadows. The observation with the highest total score is selected for use in the final image composite. However, we noticed that some image composites created using this approach had abrupt changes in image color within regions that should be relatively homogeneous (Figure A2). To mitigate this problem, we developed a weighted-available-pixels (WAP) method, which used the scores calculated by the BAP method to combine all available observations to determine the pixel values for use in the composited image. A detailed description of this method is provided in (Meng et al. 2022).

We implemented the WAP approach on the GEE platform and used it to produce clear-view composites with target dates centered around the peak growing season for all areas of China. The input images included all Landsat 8 OIL surface reflectance images acquired over China in 2020 that had cloud coverage <70%. Clouds and cloud shadows were masked using the Fmask algorithms (Zhu and Woodcock 2014a, 2014b). However, a few extremely cloudy regions (less than 5% of the total area of China) required 2019 and/or 2021 images to clear the residual clouds in the composites created using 2020 data.

Given the vast territory of China, there were substantial variations in the peak growing season across China. To account for such geographic variations, we divided the country into seven sub-regions. Figure 3 shows the approximate boundaries of these regions. The growth peak day and the date range of the

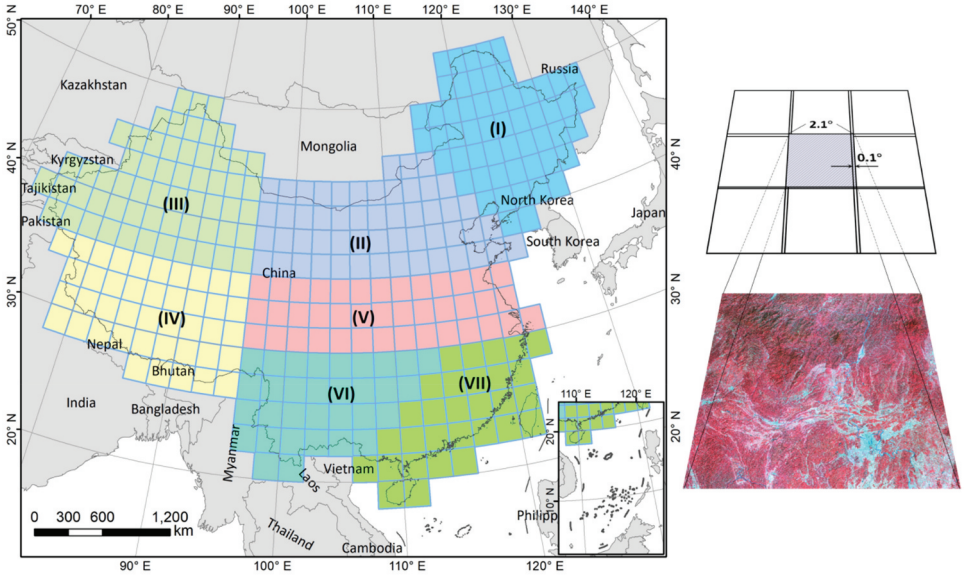

Figure 3. Spatial distribution of the 311 tiles needed to cover mainland China and the 7 sub-regions (I to VII) used in this study to account for geographic variations in the date range of growing season and growth peak days (left). Each tile covers an area of 2.1 degrees × 2.1 degrees with 0.1 degrees overlap with adjacent tiles (right).

Table 3. Growth peak day and date range of the growing season used by the WAP method for each of the 7 sub-regions.

Region ID	Location	Growth peak day	Date range of the growing season
I	Northeast	August 1st	June 15th-September15th
II	North	July 15st	June 1st-September30th
III	Northwest (north part)	July 15st	June 1st-September15th
IV	Northwest (south part)	July 15th	June 1st-September30th
V	Central	July 15th	May 15th-October15th
VI	Southwest	July 15st	May 1st-November1st
VII	South	July 1st	April 1st-November30th

growing season for each sub-region, which were required by the WAP method, are specified according to Table 3. All images acquired within the date range of the growing season were used as input to the WAP algorithm.

The output image composites had 6 spectral bands (blue, green, red, NIR, SWIR1, and SWIR2 bands) and three spectral indices (Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), and Normalized Burn Ratio (NBR)). These composites were divided into 2.1 degrees × 2.1 degrees tiles with an overlap of 0.1 degrees between adjacent tiles in geographic projection. This resulted in a total of 311 tiles needed to provide complete coverage of mainland China (Figure 3).

Initial land cover classification

The initial classification was generated using the Random Forest (RF) algorithm (Breiman 2001). Many studies have demonstrated the robustness of this algorithm for mapping land cover at regional, national, and global scales (Nguyen et al. 2020; Sales et al. 2021; Ebrahimy et al. 2021). This algorithm does not appear to be overly sensitive to mislabeled training data or other low-level random noise (Mellor et al. 2015; Pelletier et al. 2017) and can produce stable results even when the training sample size is reduced by 40%, or when up to 20% of the training samples had land cover change (Gong et al. 2019).

The training samples required by this algorithm were derived based on the land cover products described in Table 1. Because those products were developed using different classification systems with classes defined using different criteria, we used the classification scheme of the GlobeLand30 product suite in this study. The more detailed land cover types of the ChinaCover and GLC_FCS30 products were aggregated to create these 10 classes following the rules defined in Table A1. Together with the GlobeLand30 product suite, this resulted in a total of 9 land cover maps (see Table A1) that had the same 10-class classification scheme. It should be noted that because forest training samples were selected from the intersection areas of the 9 land cover products, the canopy cover threshold value for the forest class was determined by the product that had the strictest threshold, which was 20%. Therefore, the forest class in this study includes forest land dominated by trees (including arbors and sparse woods) with a canopy cover of at least 20%.

The land cover maps were then reprojected and cut into tiles to match the 311 tiles of the Landsat image composites. For each tile, we produced two maps based on the 9 land cover products. One was a union map for the forest class. In this map, a pixel had a forest class label if at least one of the 9 land cover products classified that pixel as forest. This union map was used during the post-processing stage to define the maximum extent of forest distribution in the final CAF-LC30 2020 product. The other was an intersection map, which included pixels where all 9 land cover products had the same class label at each pixel location. These pixels were then evaluated using the two reference datasets, which demonstrated that the class labels of these pixels were highly reliable. Therefore, the samples needed to train the RF algorithm were constructed based on the intersection map.

For each tile, a maximum of 10,000 random samples per class were selected from the intersection map. These samples were used to train local RF models for each tile. Previous studies have demonstrated that the use of many local classification models often results in better map products than using a single global classifier for large area land cover mapping (Gong et al. 2013; Zhang and Roy 2017; Zhang et al. 2020). The training sample sizes used in this study were determined heuristically based on the literature. Foody (2009) found that classification accuracy was positively correlated with training sample size, but it saturated after the training sample size exceeded a certain threshold. Zhu et al. (2016) proposed that the optimal training number for a Landsat scene (185 km × 185 km) would be ~20,000 pixels per

class. In this study, we found that a training size of 10,000 samples per class was sufficient to train a RF model for each tile. For tiles that did not have enough training data (i.e. <10,000 pixels for a class), the training samples were selected from the 8 adjacent tiles surrounding the target tile.

For each tile, we built five RF models. Each model had 500 decision trees trained using 70% of the samples randomly selected from the training dataset developed for that tile, and was evaluated using the remaining 30% samples. The RF model that produced the best results was used to produce the initial classification map for that tile.

Post-processing and final map generation

Because our goal was to produce an improved forest cover map, we mainly focused on improving the forest class by majority voting strategy using the given land cover products during post-classification processing (Figure 4). Based on extensive visual assessments of the initial RF classification maps, we noticed that some pixels (5.26%) in regions dominated by nonforest types such as cropland, grassland, or shrubland were classified as forest (commission error). On the other hand, 11.73% of pixels in mostly forested areas were misclassified as nonforest (omission error). The union and intersection maps derived previously based on the 9 input land cover maps were used to reduce these errors.

The maximum extent of the forest could be defined using the union map for the forest class. Very few pixels (less than 5%) located outside the union areas were true forest pixels. Therefore, (1) pixels located within the intersection map for the forest class were labeled as a forest regardless of the outcome of the RF classifications; (2) For pixels located outside the intersection region but within the union region, their land cover types were determined based on a majority voting by four land cover layers that were closest to 2020 (i.e. GlobeLand30-2020, GLC_FCS30-2020, ChinaCover-2015, and the initial RF classification map). For pixels that did not have a clear majority, the initial RF classification results were used; (3) Pixels located outside the union region were relabeled based on a majority voting by the same four land cover layers. If there was no majority over a pixel location, the class label from the RF classifications was selected. However, if that pixel was mapped as forest by the RF algorithm, it would then be reclassified according to the GlobeLand30-2020, which appeared to be more reliable than the other two product suites in general.

Product assessment

The final CAF-LC30 2020 product was evaluated using the NE-China and CTERN reference datasets to derive pixel level agreement metrics. At the provincial level, we compared forest areas calculated

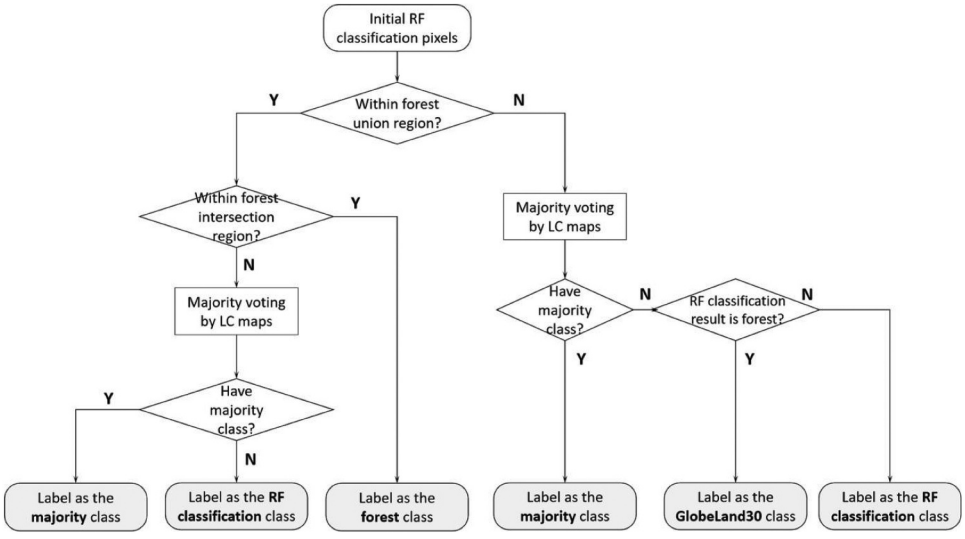

Figure 4. Flowchart of the post-processing rules used to produce the final CAF-LC30 2020 product. Only the four land cover layers that were closest to 2020 (i.e. GlobeLand30-2020, GLC_FCS30-2020, ChinaCover-2015, and the RF classification map) were used in the majority voting process.

using this product against estimates reported by the NFI program. For comparison, these assessments were also applied to four existing land cover products developed for 2020, including the GlobeLand30-2020 and GLC_FCS30-2020 that were used as inputs in this study, and the CLCD-2020 and ESA10-2020 that became available toward the end of this study. The pixel level assessment was conducted at the 30 m resolution. The ESA10-2020 dataset was resampled from 10 m to 30 m using a simple majority method.

At the pixel level, we calculated a confusion matrix for each product using each of the two reference datasets. From the confusion matrix, a suite of agreement measures, including the overall accuracy (OA), class-specific user's accuracy (UA), and producer's accuracy (PA), were calculated. Because the plots in the reference data used in this study were not selected following a probability-based sampling method (Stehman 1999, 2000), we did not know the inclusion probability of individual plots in the two reference datasets, and hence assumed that in each dataset, all plots had the same inclusion probability when calculating the accuracy measures using standard accuracy estimation equations (Congalton 1991; Olofsson et al. 2014). Therefore, the accuracy values reported in this study should serve as indicators of the agreement levels between the land cover products and available reference data, not the "true accuracies" that would be estimated using reference samples selected following a probability-based sampling method.

At the provincial level, we calculated the percentage of forest cover for each of the 31 PARMs in mainland China using the CAF-LC30 2020 and the four existing land cover products for 2020. These forest cover estimates were then compared with those reported by the 9th NFI (Table A2).

Results

Differences and complementarity of existing land cover products

Overall, the 9 maps from the three sets of input land cover products used in this study provided relatively consistent estimates of forest cover over China, which ranged from 20.24% to 22.86% (Figure 5). But they had more variability regarding the areal proportions for cropland (18.01%-25.48%), grassland (22.21%-29.88%), bareland (17.49%-30.26%), and shrubland (0.91%-7.19%). The total fraction of all other land cover types was minimal (5% or less). While the ChinaCover product suite was relatively stable between 2000 and 2015, the percent cover estimates provided by the GlobeLand30 and GLC_FCS30

Figure 5. *China*'s land cover proportions calculated using GlobeLand30 (a-c), GLC_FCS30 (d-f), and ChinaCover (g-i) for different years.

product suites were more variable over time for some land cover types. For example, the percentage of bareland reported by GlobeLand30 decreased from 30.26% in 2000 to 21.69% in 2010, while the percentage of grassland increased from 22.21% to 29.82% for the same period. The GLC_FCS30 reported a decrease in cropland from 25.48% to 20.60% and an increase in grassland from 26.64% to 29.88%.

The intersection map derived using the 9 input land cover products revealed that these products agreed with one another over 39% of the areas in mainland China (Figure 6). These areas were mostly forest (11.4%), cropland (9.2%), grassland (7.3%), and bareland (11.4%). For several classes, most of the pixels in the intersection areas were classified correctly by the intersection map. In northeast China, 99% of the plots in the NE-China reference dataset that were located within the intersection area for the cropland class were cropland plots. That agreement value was 92%, 88%, and 93% for the forest, grassland, and impervious classes, respectively. Across China, 97% of the CTERN plots located within the intersection areas for the forest class were true forest plots. Because wetlands were more difficult to classify in general, pixels in the intersection areas for this class had lower agreement values with the two reference datasets.

We defined a union area that included all pixels that were classified as forest by at least one of the 9 products and used this union mask to reduce confusion between forest and other land cover types. In northeast China, 95% of the forest plots in the NE-China dataset were located within the union mask. Across China, the union areas included 98% of the forest plots in the CTERN dataset. The fact that only 2%-5% of the forest plots in the two reference datasets were located outside the union area mask was an indicator of the usefulness of this mask for constraining the geographic extent of forest during the post-classification processing. Omission errors that may arise from not mapping forests outside the union area should be very low.

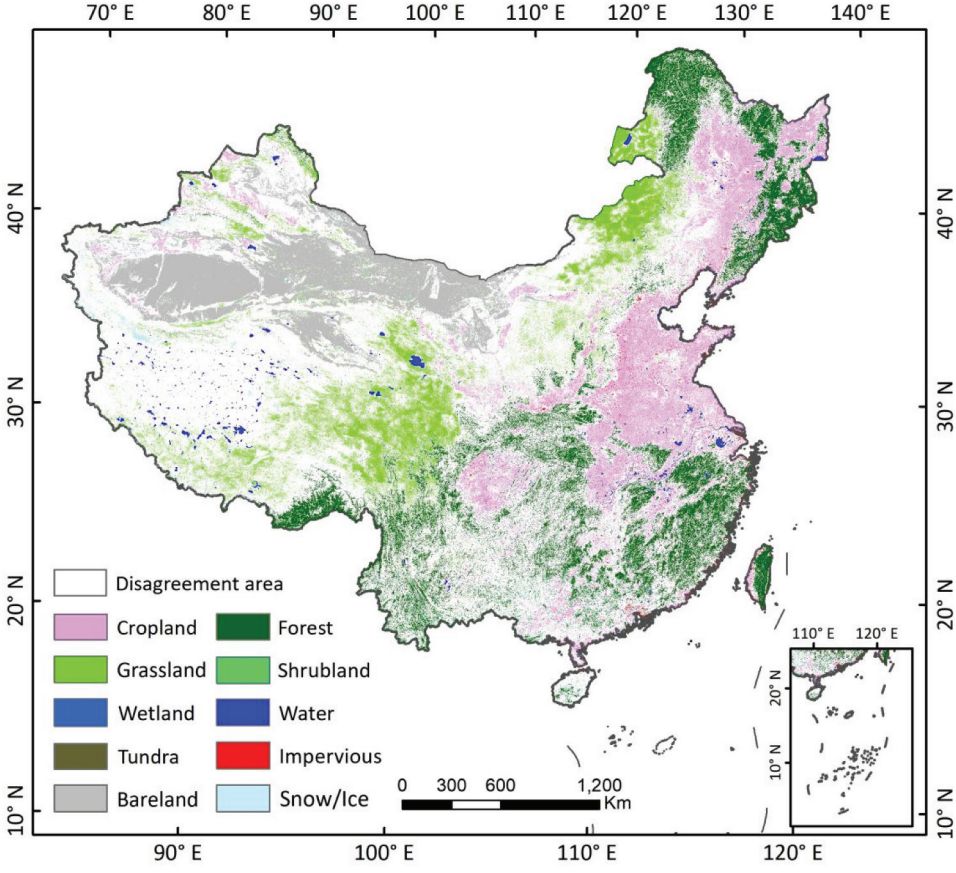

Figure 6. The intersection areas among 9 input land cover products of ten types within China.

Improvements of the CAF-LC30 2020 product over existing land cover maps

With the mapping strategy described in the method section, we produced a forest cover map for China for 2020 (Figure 7). Assessments at both pixel and provincial levels revealed that this product was better than four existing products developed for the same year over China, including GlobeLand30-2020, GLC_FCS30-2020, CLCD-2020, and ESA10-2020.

Pixel level improvements

In northeast China, the CAF-LC30 2020 map had an OA of 79.85 ± 0.83% at the 95% confidence interval when compared with field data collected over 8,895 plot locations distributed across that region (Table 4). It was significantly better than the OA values of the other four land cover products at the 95% confidence interval. The OA values of those four products were 73.47 ± 0.92% (GlobeLand30-2020), 76.91 ± 0.88% (GLC_FCS30-2020), 69.57 ± 0.96% (CLCD-2020), and 70.58 ± 0.95% (ESA10-2020). Further, the CAF-LC30 2020 map had the highest or second highest UA and PA for most of the classes covered by the reference dataset.

Throughout China, the CAF-LC30 2020 map also had a better OA with field observations over the 2,813 CTERN plots distributed across China (Table 5). While all five products had near-perfect user's accuracies (98%-99%) with CTERN data over the 2,294 forested plots, the producer's agreement of the CAF-LC30 2020 map for the forest class was 2%-4% higher than those of the other four existing products. Except for the CLCD-2020 product, the differences were statistically significant for the other three products at the 95% confidence interval. The confusion matrices of the five products derived using each of the two reference datasets are provided in Tables A3-A12.

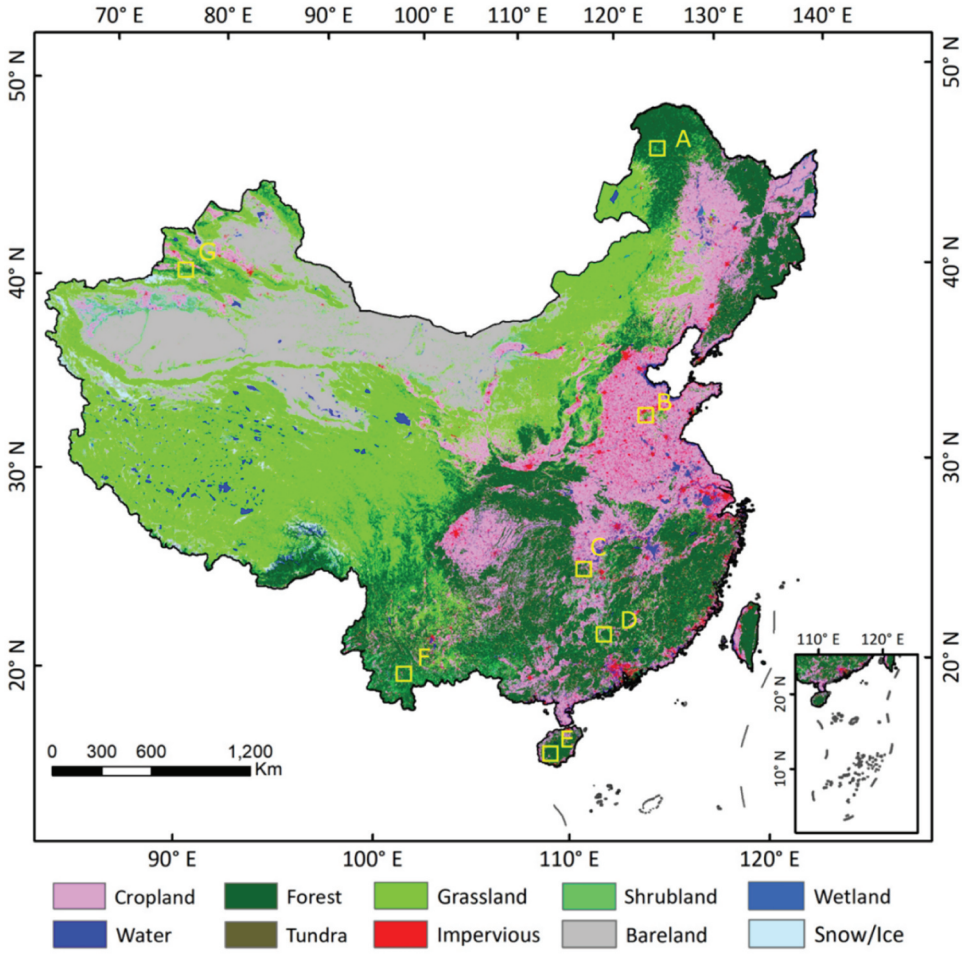

Figure 7. An overview of the final CAF-LC30 2020 map. The yellow squares indicate the approximate location of 7 selected sites where a detailed comparison of this product with four other products is presented in Figure 8.

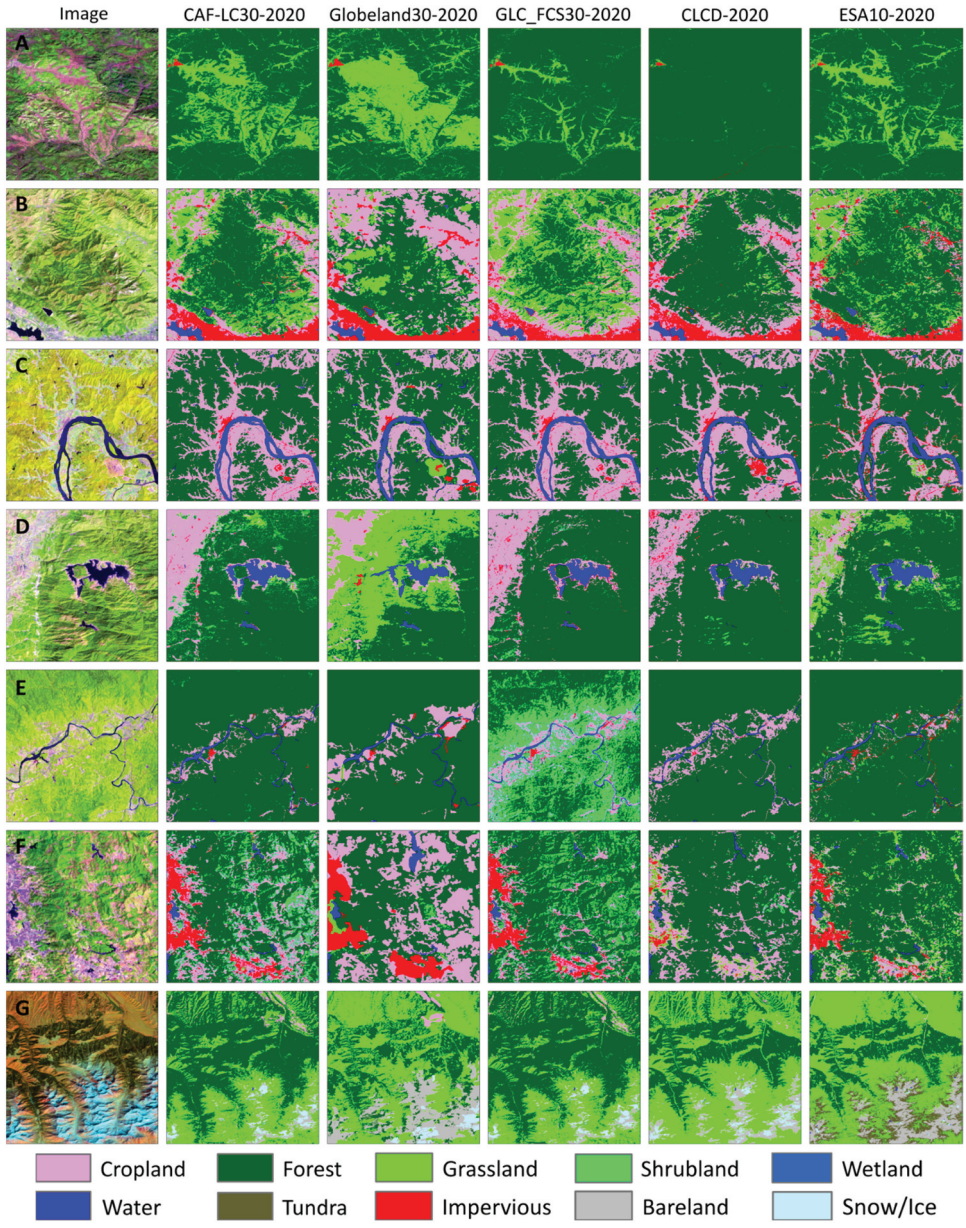

Figure 8. Comparison of the CAF-LC30 2020 map (2nd column) and four existing land cover products (3rd to 6th column) with the clear view Landsat composites (first column, Landsat 8 bands 6, 5, 4 shown in red, green, and blue) over 7 sites representing boreal forest (A), temperate continental forest (B and G), subtropical humid forest (C, D, and F), and tropical forest (E) in China. The approximate locations of the 7 sites are shown in Figure 7.

A detailed examination of the map products over different forest regions across China revealed that each existing product had different issues in different regions (Figure 8). The GlobeLand30-2020 product misclassified forest as grassland in several sites (Figure 8A, 8D, and 8F), with the ESA10-2020 product, and to a lesser degree, the CLCD-2020 product, having a similar problem. Compared with other products, the CLCD-2020 dataset overestimated forest at the expense of missing details for some nonforest classes (Figure 8A, 8B, 8D, and F)). A major issue with the GLC_FCS30-2020 product was its overestimation of shrubland (Figure 8E), especially in southern China (Figure A1 (d-f)). Derived based on clear view Landsat composites and multiple existing land cover products, the CAF-LC30 2020 map appeared to be able to provide more realistic representation of the land cover details at each of the 7 sites examined in Figure 8.

Provincial level improvements

Some of the issues with the existing land cover products were more pronounced when they were used to

Table 4. Pixel level agreements between 2020 land cover products and field survey data collected in northeast China and their ranges at the 95% confidence interval.

Product	Accuracy	Cropland	Forest	Grassland	Wetland	Water	Impervious	Bareland
CAF-LC30 2020	PA (%)	94.51 ±0.70	67.23 ±1.89	82.28 ±3.44	28.08 ±4.71	56.89 ±5.26	78.76 ±2.36	22.37 ±9.37
	UA (%)	80.14 ±1.12	88.83 ±1.46	71.30 ±3.79	74.81 ±7.43	80.83 ±4.98	70.97 ±2.48	50.00 ±16.81
				OA (%) = 79.85 ± 0.83				
GlobeLand30-2020	PA (%)	90.10 ±0.91	53.89 ±2.00	83.76 ±3.32	22.92 ±4.41	44.28 ±5.27	77.72 ±2.40	18.42 ±8.72
	UA (%)	81.40 ±1.13	87.38 ±1.70	38.47 ±2.97	68.38 ±8.43	79.89 ±5.71	60.85 ±2.49	26.92 ±12.06
				OA (%) = 73.47 ± 0.92				
GLC_FCS30-2020	PA (%)	94.00 ±0.73	58.52 ±1.98	77.22 ±3.78	22.92 ±4.41	56.89 ±5.26	79.79 ±2.31	17.11 ±8.47
	UA (%)	78.91 ±1.14	88.71 ±1.57	46.45 ±3.48	56.74 ±8.18	86.22 ±4.50	74.46 ±2.43	54.17 ±19.93
				OA (%) = 76.91 ± 0.88				
CLCD-2020	PA (%)	89.17 ±0.95	54.86 ±2.00	49.79 ±4.50	18.05 ±4.04	30.79 ±4.90	68.83 ±2.67	11.84 ±7.26
	UA (%)	77.94 ±1.18	72.08 ±2.07	36.53 ±3.71	43.75 ±8.10	100.00 ±0.00	54.55 ±2.55	56.25 ±24.31
				OA (%) = 69.57 ± 0.96				
ESA10-2020	PA (%)	86.48 ±1.04	58.73 ±1.98	69.83 ±4.13	19.77 ±4.18	48.97 ±5.31	63.82 ±2.77	17.11 ±8.47
	UA (%)	80.59 ±1.17	78.25 ±1.91	37.66 ±3.20	80.23 ±8.42	76.61 ±5.62	65.63 ±2.77	3.41 ±1.82
				OA (%) = 70.58 ± 0.95				

Table 5. Pixel level agreements between 2020 land cover products and field data collected at CTERN plots distributed across China and their ranges at the 95% confidence interval.

Product	Accuracy	Cropland	Forest	Grassland	Wetland	Water	Impervious	Bareland
CAF-LC30 2020	PA (%)	80.00 ±35.06	87.14 ±1.37	91.43 ±6.56	35.29 ±9.27	47.62 ±5.52	95.65 ±8.34	75.00 ±42.44
	UA (%)	2.09 ±2.03	99.06 ±0.42	23.88 ±5.10	64.29 ±12.55	99.34 ±1.29	28.21 ± 9.99	5.88 ±6.46
				OA (%) = 80.98 ± 0.82				
GlobeLand30-2020	PA (%)	80.00 ±35.06	82.87 ±1.54	88.57 ±7.45	35.29 ±9.27	41.27 ±5.44	100.00 ±0.00	50.00 ±49.00
	UA (%)	1.98 ±1.92	99.01 ±0.44	16.40 ±3.73	59.02 ±12.34	98.48 ±2.09	28.40 ±9.82	5.13 ±6.92
				OA (%) = 76.72 ± 0.88				
GLC_FCS30-2020	PA (%)	60.00 ±42.94	84.13 ±1.50	88.57 ±7.45	34.31 ±9.21	30.48 ±5.08	100.00 ±0.00	50.00 ±49.00
	UA (%)	1.16 ±1.30	98.92 ±0.46	20.39 ±4.53	47.30 ±11.38	96.97 ±3.38	27.06 ±9.44	4.88 ±6.59
				OA (%) = 76.47 ± 0.88				
CLCD-2020	PA (%)	60.00 ±42.94	85.31 ±1.45	71.43 ±10.58	58.82 ±9.55	40.32 ±5.42	91.30 ±11.52	25.00 ±42.44
	UA (%)	1.28 ±1.44	98.54 ±0.53	36.50 ±8.06	25.64 ±5.59	93.38 ±4.18	100.00 ±0.00	1.54 ±2.99
				OA (%) = 78.88 ± 0.85				
ESA10-2020	PA (%)	80.00 ±35.06	82.87 ±1.54	85.71 ±8.20	7.84 ±5.22	16.19 ±4.07	95.65 ±8.34	50.00 ±49.00
	UA (%)	1.59 ±1.55	98.55 ±0.53	15.31 ±3.56	33.33 ±18.86	94.44 ±6.11	16.30 ±6.23	7.14 ±9.54
				OA (%) = 72.80 ± 0.92				

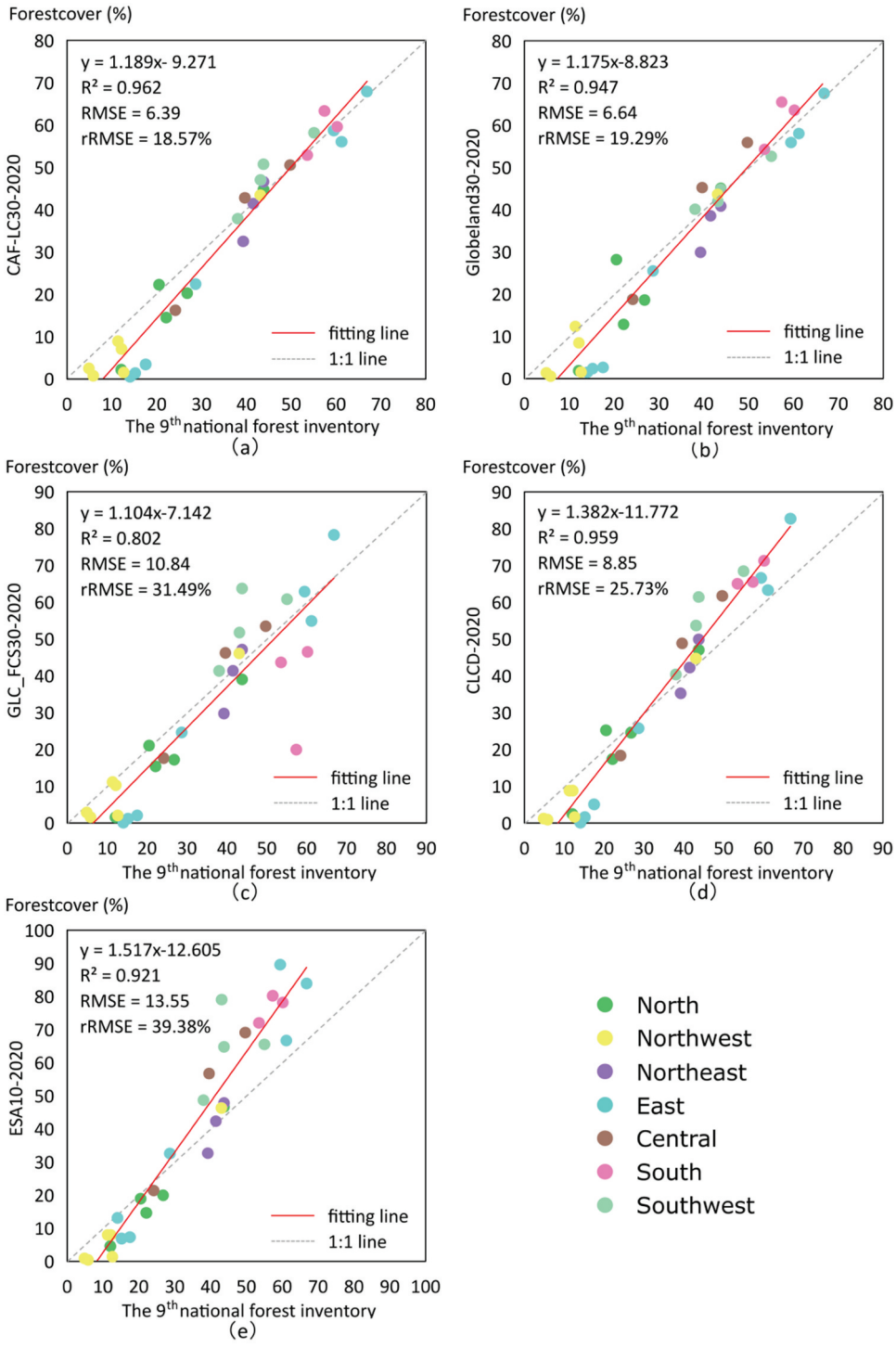

Figure 9. Comparison of the percentage of forest cover reported by the 9th NFI for each of the 31 PARMs in mainland China with those calculated based on (a) CAF-LC30 2020, (b) GlobeLand30-2020, (c) GLC_FCS30-2020, (d) CLCD-2020, and (e) ESA10-2020. Each dot represents one PARM, its color denotes the sub-regions of that PARM in China (Figure 10).

calculate forest areas at the provincial level. The percentages of forest cover calculated using the CAF-LC30 2020 for the 31 PARMs in mainland China were better correlated (R^2 = 0.962, rRMSE = 18.57%) with those reported by the 9th NFI survey conducted from 2014 to 2018 than the percentages calculated using the other four existing land cover products (Figure 9). While the GlobeLand30-2020 underestimated forest over several of the sites examined in Figure 8, those local issues did not have much impact on the provincial level forest percentage estimates derived using this product. The relationships between those estimates

and the 9th NFI survey data were only marginally worse than the relationships between estimates derived from the CAF-LC30 2020 product and the NFI data. It should be noted that while forest area estimates derived using all 5 land cover datasets for 2020 were correlated with NFI data with high R^2 values in general (Figure 9), the linear fit between the estimates from each product and NFI data was statistically different from the 1:1 line (Table A13), meaning differences between those estimates and the NFI data were statistically significant. The estimates calculated using the CAF-LC30 product had the lowest root mean square error (RMSE) when compared to the NFI data (Figure 9), indicating that this product could provide more accurate forest cover estimates at the PARM level than the other four products.

A comparison of the forest cover estimates derived from different land cover products against NFI data for individuals, PARMs revealed that each product had issues that exhibited certain regional patterns (Figure 10). For example, the ESA10-2020 product reported substantially higher forest percentages than NFI and other land cover products for several PARMs, including Chongqing, Sichuan, Guangdong, Guangxi, Hainan, Zhejiang, Hubei, and Hunan, most of which had 50% or more forest cover. On the contrary, the GLC_FCS30-2020 product mapped some forestland as shrubland or grassland in three provinces in southern China, including Guangdong, Guangxi, and Hainan. It is worth noting that the forest cover in any of the five land cover products in 2020 was underestimated when compared with the NFI data for several PARMs in northwestern China, including Qinghai, Ningxia, and Xinjiang in the northwest, Shanghai, Jiangsu, and Shandong in the east, and Tianjin in the north. This was likely due to the use of a more inclusive forest definition by the NFI program in certain situations. For example, the shrubbery or woody shrubs (cover ≥ 20% and height ≥3 m) that play significant ecological protection roles in arid or semiarid areas are considered "special forests" by the NFI survey. In arid or semiarid regions like northwestern China where Qinghai, Ningxia, and Xinjiang were located, shrubland areas with a canopy cover > 30% were classified as forestland because they were

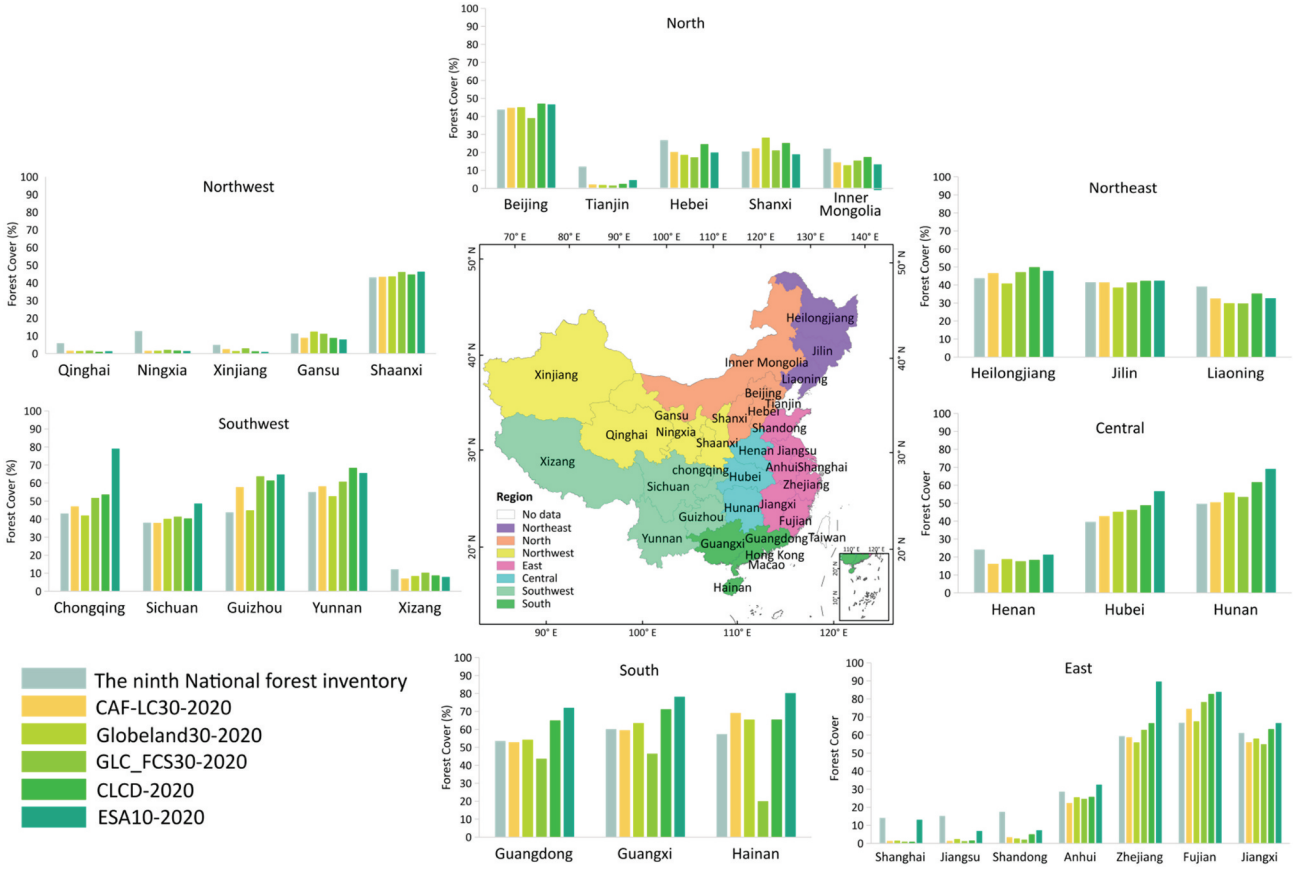

Figure 10. Comparison of provincial level forest cover estimates reported by the 9th NFI survey (1st column of each sub-graph) with those calculated based on five land cover products (2nd to 6th column of each sub-graph) over the seven sub-regions in mainland China.

important for conservation purposes. Shrub-like crops that had important economic values and had a canopy cover > 30% were also classified as forests by NFI. For PARMs like Shanghai, Jiangsu, Shandong, and Tianjin that had extremely low forest cover, classifying the shrub crops as forestland likely was the main reason why NFI reported more forestland than mapped by any of the five land cover products.

Discussion

Opportunities in a product-rich era

Free access to global 30 m or finer resolution satellite datasets from powerful cloud computing systems that are accessible at affordable or no costs greatly improved the feasibility to produce land cover products needed to capture the sub-hectare spatial details characteristic of many land cover objects, especially those related to human activities. Now that several global fine resolution (i.e. 10 m to 30 m) land cover products have been published (Chen et al. 2015; Zhang et al. 2019; Gong et al. 2019; Zanaga et al. 2021; Karra et al. 2021; Potapov et al. 2022; Friedl et al. 2022), most land areas of the globe are already covered by multiple products. Many regions have even more products available because they were also covered by country- or region-specific land cover products (Zhang et al. 2017; Manakos et al. 2018). This study demonstrated that multiple existing land cover products could be harmonized to improve forest cover mapping at the 30 m resolution.

Developed by three different research groups, the nine input land cover products used in this research did not have a high level of agreement. At the classification level considered in this study, these products agreed with one another on the land cover types for less than 40% of the pixels (Figure 6). However, the class labels of these pixels appeared to be highly accurate (> 90% of the time) for several classes, including forest, cropland, and impervious surface. Therefore, these pixels could provide high-quality training data needed to calibrate machine learning algorithms in an automated land cover mapping workflow. While land cover maps generated by machine learning algorithms typically have considerable confusions between many land cover types due to a lack of adequate separability between those types (Schneider, Friedl, and Potere 2010; Loosvelt et al. 2012), the complementarity of the input land cover products could be leveraged to reduce some of those confusions. On the one hand, since pixels in the intersection areas where the input products agreed with one another were already classified accurately by the existing land cover products, errors in classification maps produced by a machine learning algorithm over these areas could be resolved based on the input land cover products. On the other hand, the union areas for a class that included all pixels where at least one of the input products classified those pixels as that class defined the maximum extent for that class reasonably well. According to the reference datasets used in this study, the majority of the forest pixels (95% to 98%) in China were located within the union areas defined by the 9 input land cover products for the forest class. Therefore, this union area mask could be used to correct the misclassification of nonforest pixels as forest outside the mask area. Potential omission errors for the forest class that may arise from this correction should be low (likely < 5% in this study).

Advantages of the mapping approach and future improvements

The forest cover mapping approach developed in this study was designed to leverage both the intersection and union areas of the 9 input land cover products. The fact that the resultant CAF-LC30 2020 map was substantially more accurate than four existing land cover products for the year 2020 demonstrated the effectiveness of this mapping approach. Our approach greatly reduced the classification errors that the input land cover products had in different regions. For example, the GLC_FCS30 product suite mapped more shrubs and fewer forests in south and southwest China than the other two input product suites. In northeast China, the ChinaCover product suite had substantially more wetlands. Large disagreements also existed among the three product suites in arid and semiarid northwest China (Figure A1). In addition to the agreement metrics reported in the results, comprehensive visual assessments of the CAF-LC30 2020 map revealed that it provided a more accurate representation of the land cover types in these regions and many other regions where the existing land cover products had obvious errors.

Because our approach does not require region specific land cover knowledge provided by local

experts, it can be automated and used to improve land cover mapping for any region of the globe, especially for countries that have one or more region specific land cover products in addition to the available global land cover products. To what degree improvements can be achieved using this approach will be determined in part by the difference and complementarity levels of the input land cover products. In general, products developed by different producers likely will result in more improvements than products developed by the same producers, because the former tend to be more independent of one another and therefore can better complement one another toward improving land cover mapping. Caution should be taken when the input land cover products had high levels of similarities or were developed by the same producers. The intersection and union areas calculated using such products may not as valuable as discussed earlier for improving land cover mapping.

Our approach consisted of a WAP image compositing method, which was an improvement over the BAP method developed by White et al. (2014). The WAP method produced clear view composites that looked smooth and natural in areas where composites generated by the BAP method were blotchy (Figure A2). While at each location we only produced one clear view composite for a peak growing season date in this study, the WAP method can be used to produce composites for other seasons. Further improvements to forest cover mapping might be achievable by using composites created for more than one season, as the seasonal information provided by such composites may improve the separability between forest and other vegetation types (Homer et al. 2004; Huang, Homer, and Yang 2003). Further, other ancillary datasets can also be used to improve land cover mapping. For example, the OpenStreet dataset is often used to improve the mapping of urban impervious areas (Fan, Wu, and Wang 2019; Grippa et al. 2018; Zhong et al. 2020). Since 2018, the GEDI and ICESat-2 missions have collected dense samples of LiDAR measurements across the globe (Dubayah et al. 2020; Markus et al. 2017). Because LiDAR can provide highly accurate measurement of vegetation height structure (Lefsky et al. 1999; Drake et al. 2002; Bater et al. 2011; Lee et al. 2011; Potapov et al. 2021), these samples can provide high quality training data for improving forest cover mapping.

Conclusion

An approach was developed for integrating existing LC products with clear view Landsat composites to improve forest cover mapping at sub-hectare spatial resolutions. Assessments using two independent reference datasets revealed that the CAF-LC30 2020 product derived using this approach over China was more accurate than four existing land cover products, including two that were not used as input to the mapping algorithm. Its OA with field observations was 2.94% to 10.28% higher than those of the four existing land cover products in northeast China and was 2.10% to 8.18% better across China. Comprehensive visual assessments of the CAF-LC30 2020 map revealed that it provided a more accurate representation of the land cover types in regions where the existing land cover products had large classification errors. Estimates of forest cover calculated using the CAF-LC30 2020 for the 31 provinces and autonomous regions and municipalities (PARM) in mainland China were better correlated with data reported by the most recent NFI survey than those calculated using the other four existing land cover products. Therefore, the CAF-LC30 2020 product should be a better alternative for understanding China's forests in 2020 than the other four existing land cover products.

With rapid progress being made in the derivation of fine resolution land cover products at national to global scales, soon there likely will be no shortage of land cover products for many regions. However, this does not mean that the available products will have sufficient quality to meet the requirements of specific applications. This study demonstrated that improved products could be derived by leveraging those existing products. Our approach provides a framework for integrating existing land cover products, newly available remote sensing observations, and/or other ancillary data sources toward improving land cover mapping over large regions.

Acknowledgements

The CTERN plots were provided by the National Forestry and Grassland Science Data Center, the 2020 fieldwork data were provided by Prof. Zongming Wang at the Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences. Five graduate students, including Kaiyuan Shi, Yilin Fan, Ming Yan, Xiangnan Sun, and Xiaodong Niu, helped in data preparation. This work was supported by the Research Funds for the

Central Non-profit Research Institution of the Chinese Academy of Forestry (No. CAFYBB2016ZD004, CAFYBB2020ZD002, and CAFYBB2018GB001), National Key Research and Development Program of China (No. 2019YFE0126700), and the Asia-Pacific Network for Sustainable Forest Management and Rehabilitation (No. 2018P1-CAF).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Research Funds for the Central Non-profit Research Institution of the Chinese Academy of Forestry [No. CAFYBB2016ZD004], the Research Funds for the Central Non-profit Research Institution of the Chinese Academy of Forestry [No. CAFYBB2020ZD002], the Research Funds for the Central Non-profit Research Institution of the Chinese Academy of Forestry [No. CAFYBB2018GB001], National Key Research and Development Program of China [No. 2019YFE0126700], and The Asia-Pacific Network for Sustainable Forest Management and Rehabilitation [No. 2018P1-CAF].

ORCID

Yong Pang http://orcid.org/0000-0002-9760-6580

Data availability

The CAF-LC30 2020 product is available at https://www.forestdata.cn/data-detail-en.html?id=1205__TB_MET_GRID, which hosted by the National Forestry and Grassland Data Center, National Science & Technology Infrastructure of China.

References

Andrew, M. E., M. A. Wulder, and T. A. Nelson. 2014. "Potential Contributions of Remote Sensing to Ecosystem Service Assessments." *Progress in Physical Geography* 38 (3): 328–353. doi:10.1177/0309133314528942.

Arino, O., P. Bicheron, F. Achard, J. Latham, R. Witt, and J. L. Weber. 2008. "GLOBCOVER: The Most Detailed Portrait of Earth." *Esa Bulletin.bulletin Ase.European Space Agency* 136: 24–31.

Arino, O., D. Gross, F. Ranera, M. Leroy, P. Bicheron, C. Brockman, P. Defourny, et al. 2007. "GlobCover: ESA Service for Global Land Cover from MERIS." *IEEE international geoscience and remote sensing symposium*, 23-28 July, Barcelona, Spain, 2412–2415.

Arsanjani, J. J., A. Tayyebi, and E. Vaz. 2016. "GlobeLand30 as an Alternative fine-scale Global Land Cover Map: Challenges, Possibilities, and Implications for Developing Countries." *Habitat International* 55: 25–31. doi:10.1016/j.habitatint.2016.02.003.

Bartholome, E., and A. S. Belward. 2005. "GLC2000: A New Approach to Global Land Cover Mapping from Earth Observation Data." *International Journal of Remote Sensing* 26 (9): 1959–1977. doi:10.1080/01431160412331291297.

Bater, C. W., M. A. Wulder, N. C. Coops, R. F. Nelson, T. Hilker, and E. Nasset. 2011. "Stability of sample-based scanning-LiDAR-derived Vegetation Metrics for Forest Monitoring." *IEEE Transactions on Geoscience and Remote Sensing* 49 (6): 2385–2392. doi:10.1109/TGRS.2010.2099232.

Bicheron, P., V. Amberg, L. Bourg, D. Petit, M. Huc, B. Miras, C. Brockmann, et al. 2008. "Geolocation assessment of 300 m resolution MERIS Globcover ortho-rectified products." *Proceedings of the '2nd MERIS/(A)ATSR User Workshop'*, 22–26 September, Frascati, Italy.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. doi:10.1023/A:1010933404324.

Brovelli, M. A., I. Celino, M. E. Molinari, and V. Venkatachalam. 2015. "Land Cover Validation Game."In *FOSS4G Europe,* edited by Brovelli, M. A., Minghini, M., Negretti, M., 153-158. Como, Italy.

Brown de Colstoun, E. C., C. Huang, P. Wang, J. C. Tilton, B. Tan, J. Phillips, S. Niemczura, P. Y. Ling, and R. E. Wolfe. 2017. *Global Man-made Impervious Surface (GMIS) Dataset from Landsat*. Palisades, NY, USA: NASA Socioeconomic Data and Applications Center (SEDAC).

Chen, J., and J. Chen. 2018. "GlobeLand30: Operational Global Land Cover Mapping and big-data Analysis." *Science China. Earth Sciences* 61 (10): 1533–1534. doi:10.1007/s11430-018-9255-3.

Chen, J., J. Chen, A. Liao, X. Cao, L. Chen, X. Chen, C. He, et al. 2015. "Global Land Cover Mapping at 30 M Resolution: A POK-based Operational Approach." *ISPRS Journal of Photogrammetry and Remote Sensing* 103: 7–27. doi:10.1016/j.isprsjprs.2014.09.002.

Congalton, R. 1991. "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data." *Remote Sensing of Environment* 37: 35–46. doi:10.1016/0034-4257(91)90048-B.

DeFries, R., M. Hansen, and J. R. G. Townshend. 1995. "Global Discrimination of Land Cover Types from Metrics Derived from AVHRR Pathfinder Data." *Remote Sensing of Environment* 54 (3): 209–222. doi:10.1016/0034-4257(95)00142-5.

DeFries, R. S., M. Hansen, J. R. G. Townshend, and R. Sohlberg. 1998. "Global Land Cover Classifications at 8 Km Spatial Resolution: The Use of Training Data Derived from Landsat Imagery in Decision Tree Classifiers." *International Journal of Remote Sensing* 19 (16): 3141–3168. doi:10.1080/014311698214235.

Drake, J. B., R. O. Dubayah, R. G. Knox, D. B. Clark, and J. B. Blair. 2002. "Sensitivity of large-footprint Lidar to Canopy Structure and Biomass in a Neotropical Rainforest." *Remote*

Sensing of Environmen 81 (2–3): 378–392. doi:10.1016/S0034-4257(02)00013-5.

Dubayah, R., J. B. Blair, S. Goetz, L. Fatoyinbo, M. Hansen, S. Healey, M. Hofton, et al. 2020. "The Global Ecosystem Dynamics Investigation: High-resolution Laser Ranging of the Earth's Forests and Topography." *Science of Remote Sensing* 1: 100002. doi:10.1016/j.srs.2020.100002.

Ebrahimy, H., B. Mirbagheri, A. A. Matkan, and M. Azadbakht. 2021. "Per-pixel Land Cover Accuracy Prediction: A Random forest-based Method with Limited Reference Sample Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 172: 17–27. doi:10.1016/j.isprsjprs.2020.11.024.

Fan, W., C. Wu, and J. Wang. 2019. "Improving Impervious Surface Estimation by Using Remote Sensed Imagery Combined with Open Street Map points-of-interest (POI) Data." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 12 (11): 4265–4274. doi:10.1109/JSTARS.2019.2911525.

Feng, M., J. O. Sexton, C. Huang, A. Anand, S. Channan, X. P. Song, D. X. Song, D. H. Kim, P. Noojipady, and J. R. Townshend. 2016. "Earth Science Data Records of Global Forest Cover and Change: Assessment of Accuracy in 1990, 2000, and 2005 Epochs." *Remote Sensing of Environment* 184: 73–85. doi:10.1016/j.rse.2016.06.012.

Foley, J. A., R. DeFries, G. P. Asner, C. Barford, G. Bonan, S. R. Carpenter, F. S. Chapin, et al. 2005. "Global Consequences of Land Use." *Science* 309 (5734): 570–574. doi:10.1126/science.1111772.

Foody, G. M. 2002. "Status of Land Cover Classification Accuracy Assessment." *Remote Sensing of Environment* 80 (1): 185–201. doi:10.1016/S0034-4257(01)00295-4.

Foody, G. M. 2009. "Sample Size Determination for Image Classification Accuracy Assessment and Comparison." *International Journal of Remote Sensing* 30 (20): 5273–5291. doi:10.1080/01431160903130937.

Friedl, M. A., D. K. McIver, J. C. F. Hodges, X. Y. Zhang, C. E. Muchoney, A. H. Strahler, C. E. Woodcock, et al. 2002. "Global Land Cover Mapping from MODIS: Algorithms and Early Results." *Remote Sensing of Environment* 83 (1–2): 287–302. doi:10.1016/S0034-4257(02)00078-0.

Friedl, M. A., and D. Sulla-Menashe. 2019. "MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006." *NASA EOSDIS Land Processes DAAC*. 10:200.

Friedl, M. A., D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X. Huang. 2010. "MODIS Collection 5 Global Land Cover: Algorithm Refinements and Characterization of New Datasets." *Remote Sensing of Environment* 114: 168–182. doi:10.1016/j.rse.2009.08.016.

Friedl, M. A., C. E. Woodcock, P. Olofsson, Z. Zhu, T. Loveland, R. Stanimirova, P. Arevalo, et al. 2022. "Medium Spatial Resolution Mapping of Global Land Cover and Land Cover Change across Multiple Decades from Landsat." *Frontiers in Remote Sensing* 3. doi:10.3389/frsen.2022.894571.

Fritz, S., L. See, and F. Rembold. 2010. "Comparison of Global and Regional Land Cover Maps with Statistical Information for the Agricultural Domain in Africa." *International Journal of Remote Sensing* 31 (9): 2237–2256. doi:10.1080/01431160902946598.

Fritz, S., L. You, A. Bun, L. See, I. McCallum, C. Schill, C. Perger, J. Liu, M. Hansen, and M. Obersteiner. 2011. "Cropland for sub-Saharan Africa: A Synergistic Approach Using Five Land Cover Data Sets." *Geophysical Research Letters* 38 (4). doi:10.1029/2010GL046213.

Gong, P., H. Liu, M. Zhang, C. Li, J. Wang, H. Huang, N. Clinton, et al. 2019. "Stable Classification with Limited Sample: Transferring a 30-m Resolution Sample Set Collected in 2015 to Mapping 10-m Resolution Global Land Cover in 2017." *Science Bulletin* 64 (6): 370–373. doi:10.1016/j.scib.2019.03.002.

Gong, P., J. Wang, L. Yu, Y. C. Zhao, Y. Y. Zhao, L. Liang, Z. G. Niu, and Y. L. Chen. 2013. "Finer Resolution Observation and Monitoring of Global Land Cover: First Mapping Results with Landsat TM and ETM+ Data." *International Journal of Remote Sensing* 34 (7): 2607–2654. doi:10.1080/01431161.2012.748992.

Griffiths, P., V. D. L. Sebastian, T. Kuemmerle, and P. Hostert. 2013. "A pixel-based Landsat Compositing Algorithm for Large Area Land Cover Mapping." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6 (5): 2088–2101. doi:10.1109/JSTARS.2012.2228167.

Grippa, T., S. Georganos, S. Zarougui, P. Bognounou, E. Diboulo, Y. Forget, M. Lennert, S. Vanhuysse, N. Mboga, and E. Wolff. 2018. "Mapping Urban Land Use at Street Block Level Using Openstreetmap, Remote Sensing Data, and Spatial Metrics." *ISPRS International Journal of Geo-Information* 7 (7): 246. doi:10.3390/ijgi7070246.

Gutman, G., R. Byrnes, J. Masek, S. Covington, C. Justice, S. Franks, and R. Headley. 2008. "Towards Monitoring land-cover and land-use Changes at a Global Scale: The Global Land Survey 2005." *Photogrammetric Engineering and Remote Sensing* 74: 6–10.

Gutman, G., C. Huang, G. Chander, P. Noojipady, and J. Masek. 2013. "Assessment of the NASA-USGS Global Land Survey (GLS) Datasets." *Remote Sensing of Environment* 134: 249–265. doi:10.1016/j.rse.2013.02.026.

Hansen, M. C., R. S. DeFries, J. R. G. Townshend, and R. Sohlberg. 2000. "Global Land Cover Classification at 1 Km Spatial Resolution Using a Classification Tree Approach." *International Journal of Remote Sensing* 21 (6–7): 1331–1364. doi:10.1080/014311600210209.

Hansen, M. C., and T. R. Loveland. 2012. "A Review of Large Area Monitoring of Land Cover Change Using Landsat Data." *Remote Sensing of Environment* 122: 66–74. doi:10.1016/j.rse.2011.08.024.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, et al. 2013. "High-resolution Global Maps of 21st-century Forest Cover Change." *Science* 342 (6160): 850–853. doi:10.1126/science.1244693.

Hermosilla, T., M. A. Wulder, J. C. White, N. C. Coops, and G. W. Hobart. 2015. "An Integrated Landsat Time Series Protocol for Change Detection and Generation of Annual

gap-free Surface Reflectance Composites." *Remote Sensing of Environment* 158: 220–234. doi:10.1016/j.rse.2014.11.005.

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan. 2004. "Development of a 2001 National Land Cover Database for the United States." *Photogrammetric Engineering and Remote Sensing* 70 (7): 829–840. doi:10.14358/PERS.70.7.829.

Huang, C., C. Homer, and L. Yang. 2003. "Regional Forest Land Cover Characterization Using Medium Spatial Resolution Satellite Data." In *Methods and Applications for Remote Sensing of Forests: Concepts and Case Studies*, edited by Wulder, M. A., Franklin, S. E., 389–410. Boston: Kluwer Academic Publishers.

Jung, M., K. Henkel, M. Herold, and G. Churkina. 2006. "Exploiting Synergies of Global Land Cover Products for Carbon Cycle Modeling." *Remote Sensing of Environment* 101 (4): 534–553. doi:10.1016/j.rse.2006.01.020.

Karra, K., C. Kontgis, Z. Statman-Weil, J. C. Mazzariello, M. Mathis, and S. P. Brumby. 2021. "Global Land use/land Cover with Sentinel 2 and Deep Learning." *IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 11-16 July, Brussels, Belgium, 4704–4707.

Kim, D. H., J. O. Sexton, P. Noojipady, C. Huang, A. Anand, S. Channan, M. Feng, and J. R. Townshend. 2014. "Global, Landsat-based forest-cover Change from 1990 to 2000." *Remote Sensing of Environment* 155: 178–193. doi:10.1016/j.rse.2014.08.017.

Lee, S., W. Ni-Meister, W. Yang, and Q. Chen. 2011. "Physically Based Vertical Vegetation Structure Retrieval from ICESat Data: Validation Using LVIS in White Mountain National Forest, New Hampshire, USA." *Remote Sensing of Environment* 115 (11): 2776–2785. doi:10.1016/j.rse.2010.08.026.

Lefsky, M. A., W. B. Cohen, S. A. Acker, G. G. Parker, T. A. Spies, and D. Harding. 1999. "Lidar Remote Sensing of the Canopy Structure and Biophysical Properties of Douglas-fir Western Hemlock Forests." *Remote Sensing of Environment* 70 (3): 339–361. doi:10.1016/S0034-4257(99)00052-8.

Liu, H., P. Gong, J. Wang, X. Wang, G. Ning, and B. Xu. 2021. "Production of Global Daily Seamless Data Cubes and Quantification of Global Land Cover Change from 1985 to 2020-iMap World 1.0." *Remote Sensing of Environment* 258: 112364. doi:10.1016/j.rse.2021.112364.

Loosvelt, L., J. Peters, H. Skriver, B. D. Baets, and N. Verhoest. 2012. "Impact of Reducing Polarimetric SAR Input on the Uncertainty of Crop Classifications Based on the Random Forests Algorithm." *IEEE Transactions on Geoscience and Remote Sensing* 50 (10): 4185–4200. doi:10.1109/TGRS.2012.2189012.

Loveland, T. R., B. C. Reed, J. F. Brown, D. O. Ohlen, Z. Zhu, L. Yang, and J. W. Merchant. 2000. "Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1km AVHRR Data." *International Journal of Remote Sensing* 21 (6–7): 1303–1330. doi:10.1080/014311600210191.

Manakos, I., M. Tomaszewska, I. Gkinis, O. Brovkina, L. Filchev, L. Genc, I. Z. Gitas, et al. 2018. "Comparison of Global and Continental Land Cover Products for Selected Study Areas in South Central and Eastern European Region." *Remote Sensing* 10 (12): 1967. doi:10.3390/rs10121967.

Markus, T., T. Neumann, A. Martino, W. Abdalati, K. Brunt, B. Csatho, S. Farrell, et al. 2017. "The Ice, Cloud, and Land Elevation Satellite-2 (Icesat-2): Science Requirements, Concept, and Implementation." *Remote Sensing of Environment* 190: 260–273. doi:10.1016/j.rse.2016.12.029.

Mellor, A., S. Boukir, A. Haywood, and S. Jones. 2015. "Exploring Issues of Training Data Imbalance and Mislabelling on Random Forest Performance for Large Area Land Cover Classification Using the Ensemble Margin." *ISPRS Journal of Photogrammetry and Remote Sensing* 105: 155–168. doi:10.1016/j.isprsjprs.2015.03.014.

Meng, S. L., Y. Pang, and Z. Y. Li. 2022. "Compose cloud-free image based Landsat 8 and Sentinel 2 with improved reflectance consistency: case studies in China." *in preparation*.

Nguyen, L. H., D. R. Joshi, D. E. Clay, and G. M. Henebry. 2020. "Characterizing Land cover/land Use from Multiple Years of Landsat and MODIS Time Series: A Novel Approach Using Land Surface Phenology Modeling and Random Forest Classifier." *Remote Sensing of Environment* 238: 111017. doi:10.1016/j.rse.2018.12.016.

Olofsson, P., G. M. Foody, M. Herold, S. V. Stehman, C. E. Woodcock, and M. A. Wulder. 2014. "Good Practices for Estimating Area and Assessing Accuracy of Land Change." *Remote Sensing of Environment* 148: 42–57. doi:10.1016/j.rse.2014.02.015.

Pekel, J. F., A. Cottam, N. Gorelick, and A. S. Belward. 2016. "High-resolution Mapping of Global Surface Water and Its long-term Changes." *Nature* 540 (7633): 418–422. doi:10.1038/nature20584.

Pelletier, C., S. Valero, J. Inglada, N. Champion, G. Dedieu, and G. Dedieu. 2017. "Effect of Training Class Label Noise on Classification Performances for Land Cover Mapping with Satellite Image Time Series." *Remote Sensing* 9 (2): 173. doi:10.3390/rs9020173.

Pittman, K., M. C. Hansen, I. Becker-Reshef, P. V. Potapov, and C. O. Justice. 2010. "Estimating Global Cropland Extent with Multi-Year MODIS Data." *Remote Sensing* 2 (7): 1844–1863. doi:10.3390/rs2071844.

Potapov, P., M. C. Hansen, A. Pickens, A. Hernandez-Serna, A. Tyukavina, S. Turubanova, V. Zalles, et al. 2022. "The Global 2000-2020 Land Cover and Land Use Change Dataset Derived from the Landsat Archive: First Results." *Frontiers in Remote Sensing* 3.

Potapov, P., X. Li, A. Hernandez-Serna, A. Tyukavina, M. C. Hansen, A. Kommareddy, A. Pickens, et al. 2021. "Mapping Global Forest Canopy Height through Integration of GEDI and Landsat Data." *Remote Sensing of Environment* 253: 112165. doi:10.1016/j.rse.2020.112165.

Roy, D. P., J. Ju, K. Kline, P. L. Scaramuzza, V. Kovalskyy, M. Hansen, T. R. Loveland, E. Vermote, and C. Zhang. 2010. "Web-enabled Landsat Data (WELD): Landsat ETM+ Composited Mosaics of the Conterminous United States." *Remote Sensing of Environment* 114 (1): 35–49. doi:10.1016/j.rse.2009.08.011.

Sales, M. H., S. De Bruin, C. Souza, and M. Herold. 2021. "Land Use and Land Cover Area Estimates from Class Membership Probability of a Random Forest Classification." *IEEE*

Transactions on Geoscience and Remote Sensing 60: 1–11. doi:10.1109/TGRS.2021.3080083.

Schneider, A., M. A. Friedl, and D. Potere. 2010. "Mapping Global Urban Areas Using MODIS 500-m Data: New Methods and Datasets Based on 'Urban Ecoregions." *Remote Sensing of Environment* 114 (8): 1733–1746. doi:10.1016/j.rse.2010.03.003.

Sexton, J. O., P. Noojipady, A. Anand, X. P. Song, S. McMahon, C. Huang, M. Feng, S. Channan, and J. R. Townshend. 2015. "A Model for the Propagation of Uncertainty from Continuous Estimates of Tree Cover to Categorical Forest Cover and Change." *Remote Sensing of Environment* 156: 418–425. doi:10.1016/j.rse.2014.08.038.

Sexton, J. O., X. P. Song, M. Feng, P. Noojipady, A. Anand, C. Huang, D. H. Kim, et al. 2013. "Global, 30-m Resolution Continuous Fields of Tree Cover: Landsat-based Rescaling of MODIS Continuous Fields and lidar-based Estimates of Error." *International Journal of Digital Earth* 6 (5): 427–448. doi:10.1080/17538947.2013.786146.

Song, X. P., C. Huang, M. Feng, J. O. Sexton, S. Channan, and J. R. Townshend. 2014. "Integrating Global Land Cover Products for Improved Forest Cover Characterization: An Application in North America." *International Journal of Digital Earth* 7 (9): 709–724. doi:10.1080/17538947.2013.856959.

Song, X. P., C. Huang, and J. R. Townshend. 2017. "Improving Global Land Cover Characterization through Data Fusion." *Geo-spatial Information Science*: 1–10.

Stehman, S. V. 1999. "Basic Probability Sampling Designs for Thematic Map Accuracy Assessment." *International Journal of Remote Sensing* 20 (12): 2423–2441. doi:10.1080/014311699212100.

Stehman, S. V. 2000. "Practical Implications of design-based Sampling Inference for Thematic Map Accuracy Assessment." *Remote Sensing of Environment* 72 (1): 35–45. doi:10.1016/S0034-4257(99)00090-5.

Tateishi, R., N. T. Hoan, T. Kobayashi, B. Alsaaideh, G. Tana, and D. X. Phong. 2011. "Production of Global Land Cover data-GLCNMO." *Journal of Geography and Geology* 4 (1): 22–49.

Townshend, J. R. G. 1992. "Land Cover." *International Journal of Remote Sensing* 13 (6–7): 1319–1328. doi:10.1080/01431169208904193.

Townshend, J. R. G. 1994. "Global Data Sets for Land Applications from the Advanced Very High Resolution Radiometer: An Introduction." *International Journal of Remote Sensing* 15 (17): 3319–3333. doi:10.1080/01431169408954333.

Townshend, J. R., J. G. Masek, C. Huang, E. F. Vermote, F. Gao, S. Channan, J. Sexton, et al. 2012. "Global Characterization and Monitoring of Forest Cover Using Landsat Data: Opportunities and Challenges." *International Journal of Digital Earth* 5 (5): 373–397. doi:10.1080/17538947.2012.713190.

Tucker, C. J., D. M. Grant, and J. D. Dykstra. 2004. "NASA's Global Orthorectified Landsat Data Set." *Photogrammetric Engineering and Remote Sensing* 70 (3): 313–322. doi:10.14358/PERS.70.3.313.

White, J. C., M. A. Wulder, G. W. Hobart, J. E. Luther, T. Hermosilla, P. Griffiths, N. C. Coops, et al. 2014. "Pixel-based Image Compositing for large-area Dense Time Series Applications and Science." *Canadian Journal of Remote Sensing* 40 (3): 192–212. doi:10.1080/07038992.2014.945827.

Woodcock, C. E., R. Allen, M. Anderson, A. Belward, R. Bindschadler, W. Cohen, F. Gao, et al. 2008. "Free Access to Landsat Imagery." *Science* 320 (5879): 1011. doi:10.1126/science.320.5879.1011a.

Wulder, M. A., N. C. Coops, D. P. Roy, J. C. White, and T. Hermosilla. 2018. "Land Cover 2.0." *International Journal of Remote Sensing* 39 (12): 4254–4284. doi:10.1080/01431161.2018.1452075.

Wulder, M. A., J. G. Masek, W. B. Cohen, T. R. Loveland, and C. E. Woodcock. 2012. "Opening the Archive: How Free Data Has Enabled the Science and Monitoring Promise of Landsat." *Remote Sensing of Environment* 122 (1): 2–10. doi:10.1016/j.rse.2012.01.010.

Wulder, M. A., J. C. White, T. R. Loveland, C. E. Woodcock, A. S. Belward, W. B. Cohen, E. A. Fosnight, J. Shaw, J. G. Masek, and D. P. Roy. 2016. "The Global Landsat Archive: Status, Consolidation, and Direction." *Remote Sensing of Environment* 185: 271–283. doi:10.1016/j.rse.2015.11.032.

Wu, B., J. Qian, Y. Zeng, C. Qi, and Y. Song. 2017. *Land Cover Map of the People's Republic of China((1: 1 000 000))*. Beijing, China: Sinomap press.

Yang, J., and X. Huang. 2021. "The 30 M Annual Land Cover Dataset and Its Dynamics in China from 1990 to 2019." *Earth System Science Data* 13 (8): 3907–3925. doi:10.5194/essd-13-3907-2021.

Yang, C., Q. Y. Huang, Z. L. Li, K. Liu, and F. Hu. 2017. "Big Data and Cloud Computing: Innovation Opportunities and Challenges." *International Journal of Digital Earth* 10 (1): 13–53. doi:10.1080/17538947.2016.1239771.

Zanaga, D., R. Van De Kerchove, W. De Keersmaecker, N. Souverijns, C. Brockmann, R. Quast, J. Wevers, et al. 2021. "ESA WorldCover 10 m 2020 v100".

Zhang, X., L. Liu, X. Chen, Y. Gao, S. Xie, and J. Mi. 2021. "GLC_FCS30: Global land-cover Product with Fine Classification System at 30 M Using time-series Landsat Imagery." *Earth System Science Data* 13 (6): 2753–2776. doi:10.5194/essd-13-2753-2021.

Zhang, X., L. Liu, X. Chen, S. Xie, and Y. Gao. 2019. "Fine land-cover Mapping in China Using Landsat Datacube and an Operational SPECLib-based Approach." *Remote Sensing* 11 (9): 1056. doi:10.3390/rs11091056.

Zhang, X., L. Liu, C. Wu, X. Chen, Y. Gao, S. Xie, and B. Zhang. 2020. "Development of a Global 30 M Impervious Surface Map Using Multisource and Multitemporal Remote Sensing Datasets with the Google Earth Engine Platform." *Earth System Science Data* 12 (3): 1625–1648. doi:10.5194/essd-12-1625-2020.

Zhang, H. K., and D. P. Roy. 2017. "Using the 500 M MODIS Land Cover Product to Derive a Consistent Continental Scale 30 M Landsat Land Cover Classification." *Remote Sensing of Environment* 197: 15–34. doi:10.1016/j.rse.2017.05.024.

Zhong, Y., Y. Su, S. Wu, Z. Zheng, J. Zhao, A. Ma, Q. Zhu, et al. 2020. "Open-source data-driven Urban land-use Mapping

Integrating point-line-polygon Semantic Objects: A Case Study of Chinese Cities." *Remote Sensing of Environment* 247: 111838. doi:10.1016/j.rse.2020.111838.

Zhu, Z., A. L. Gallant, C. E. Woodcock, B. Pengra, P. Olofsson, T. R. Loveland, S. Jin, D. Dagal, L. Yang, and R. F. Auch. 2016. "Optimizing Selection of Training and Auxiliary Data for Operational Land Cover Classification for the LCMAP Initiative." *ISPRS Journal of Photogrammetry and Remote Sensing* 122: 206–221. doi:10.1016/j.isprsjprs.2016.11.004.

Zhu, Z., and C. E. Woodcock. 2014a. "Automated Cloud, Cloud Shadow, and Snow Detection in Multitemporal Landsat Data: An Algorithm Designed Specifically for Monitoring Land Cover Change." *Remote Sensing of Environment* 152: 217–234.

Zhu, Z., and C. E. Woodcock. 2014b. "Continuous Change Detection and Classification of Land Cover Using All Available Landsat Data." *Remote Sensing of Environment* 144: 152–171. doi:10.1016/j.rse.2014.01.011.

Zhu, Z., C. E. Woodcock, C. Holden, and Z. Yang. 2015. "Generating Synthetic Landsat Images Based on All Available Landsat Data: Predicting Landsat Surface Reflectance at Any Given Time." *Remote Sensing of Environment* 162: 67–83. doi:10.1016/j.rse.2015.02.009.

Zhu, Z., M. A. Wulder, D. P. Roy, C. E. Woodcock, M. C. Hansen, V. C. Radeloff, S. P. Healey, et al. 2019. "Benefits of the Free and Open Landsat Data Policy." *Remote Sensing of Environment* 224: 382–385. doi:10.1016/j.rse.2019.02.016.

Appendix

Table A1. Land cover types used in this study and their relationships to those used by the other land cover products.

Target Code	Target category	GlobeLand30 code	GlobalLand30 category	ChinaCover code	ChinaCover category	GLC_FCS30 code	GLC_FCS30 category
1	Cropland	10	Cropland	301	Paddy field	10	Rainfed cropland
				302	Dry farmland	11	Herbaceous cover cropland
						20	Irrigated cropland
2	Forest	20	Forest	101	Evergreen broadleaf forest	50	Evergreen broadleaf forest
				102	Deciduous broadleaf forest	60	Deciduous broadleaf forest
				103	Evergreen needleleaf forest	61	Open deciduous broadleaf forest (0.15< fc<0.4)
				104	Deciduous needleleaf forest	62	Closed Deciduous broadleaf forest (fc>0.4)
				105	Broadleaf and needleleaf mixed forest	70	Evergreen needleleaf forest
				109	Sparse forest	71	Open evergreen needleleaf forest (0.15< fc <0.4)
				111	Tree orchard	72	Closed evergreen needleleaf forest (fc >0.4)
				113	Tree garden	80	Deciduous needleleaf forest
						81	Open deciduous needleleaf forest (0.15< fc <0.4)
						82	Closed deciduous needleleaf forest (fc >0.4)
						90	Mixed-leaf forest
						12	Tree cover (orchard)
3	Grassland	30	Grassland	201	Temperate steppe	120	Grassland
				202	Alpine steppe	121	Sparse vegetation associated with grassland (fc<0.15)
				203	Temperate meadow	122	Sparse herbaceous cover (fc<0.15)
				204	Alpine meadow		
				205	Tussock		
				206	Sparse grassland		
				207	Lawn		
4	Shrubland	40	Shrubland	106	Evergreen broadleaf shrubland	130	Shrubland
				107	Deciduous broadleaf shrubland	140	Evergreen shrubland
				108	Evergreen needleleaf shrubland	150	Deciduous shrubland
				110	Sparse shrubland	152	Sparse Sparse vegetation associated With shrubland (fc<0.15)
				112	Shrub orchard		
				114	Shrub garden		
5	Wetland	50	Wetland	401	Tree wetland	153	Wetlands
				402	Shrub wetland		
				403	Herbaceous wetland		
6	Water	60	Water	404	Lake	180	Water
				405	Reservoir/Pond		
				406	River		
				407	Canal/Channel		
7	Tundra	70	Tundra	601	Moss/Lichen	190	Lichens and mosses
8	Impervious	80	Impervious	501	Settlement	200	Impervious surface
				502	Transportation land		
				503	Mining field		
9	Bareland	90	Bareland	602	Bare rock	201	Bare areas
				603	Gobi	202	Consolidated bare areas
				604	Bare soil	210	Unconsolidated bare areas
				605	Desert		
				606	Salina		
10	Snow/Ice	100	Snow/Ice	607	Permanent ice/snow	220	Permanent ice and snow

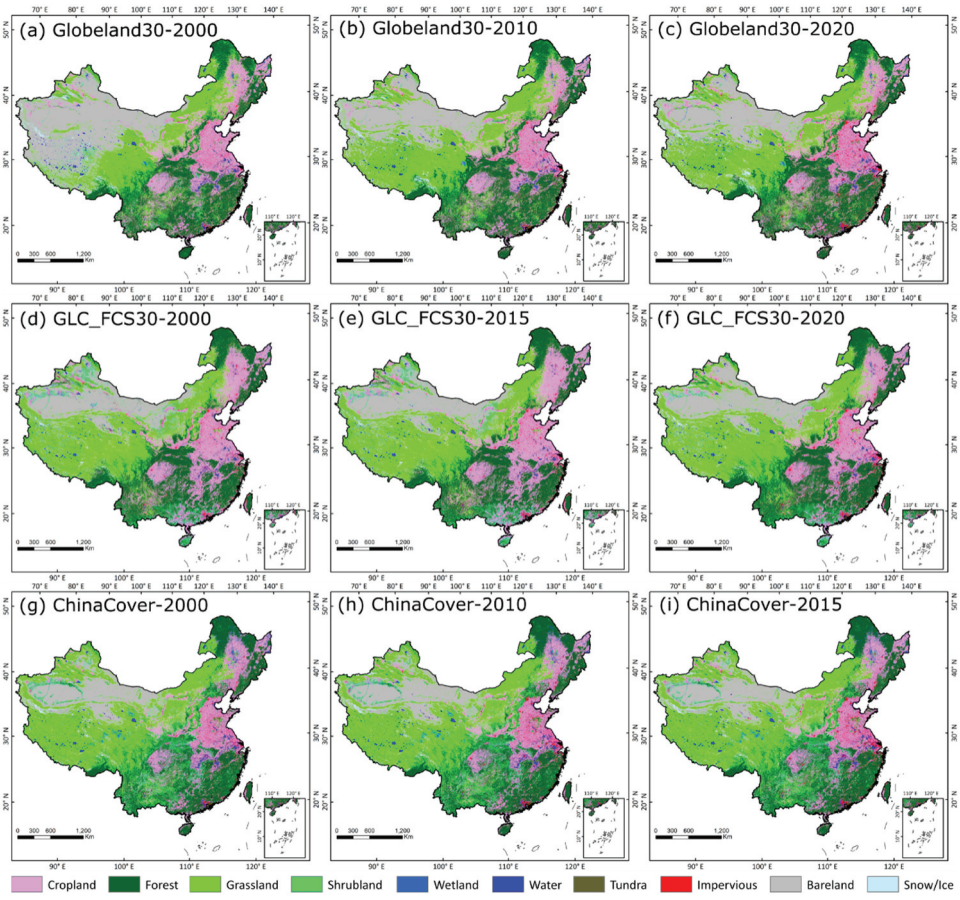

Figure A1. Overview maps of the nine input LC products after unifying their classification systems according to Table A1 to create the target LC types used in this study.

Table A2. The percentage of forest coverage reported by the 9th NFI for the 31 provinces, autonomous regions and municipalities (PARM) listed in alphabetical order).

ID	Province	Forest cover	ID	Province	Forest cover	ID	Province	Forest cover
1	Anhui	28.65	12	Henan	24.14	23	Shandong	17.51
2	Beijing	43.77	13	Hubei	39.61	24	Shanghai	14.04
3	Chongqing	43.11	14	Hunan	49.69	25	Shanxi	20.5
4	Fujian	66.8	15	Inner Mongolia	22.1	26	Sichuan	38.03
5	Gansu	11.33	16	Jiangsu	15.2	27	Tianjin	12.07
6	Guangdong	53.52	17	Jiangxi	61.16	28	Xinjiang	4.87
7	Guangxi	60.17	18	Jilin	41.49	29	Xizang	12.14
8	Guizhou	43.77	19	Liaoning	39.24	30	Yunnan	55.04
9	Hainan	57.36	20	Ningxia	12.63	31	Zhejiang	59.43
10	Hebei	26.78	21	Qinghai	5.82			
11	Heilongjiang	43.78	22	Shaanxi	43.06			

Table A3. Pixel level agreements between the CAF-LC30 2020 product and field plot data in northeast China and their ranges at the 95% confidence interval.

	Cropland	Forest	Grassland	Wetland	Water	Impervious	Bareland	Total	UA (%)
Cropland	3894	540	40	104	51	205	25	4859	80.14 ± 1.12
Forest	74	1598	10	54	31	25	7	1799	88.83 ± 1.46
Grassland	61	17	390	54	0	10	15	547	71.30 ± 3.79
Wetland	4	6	1	98	18	3	1	131	74.81 ± 7.43
Water	3	15	1	13	194	3	11	240	80.83 ± 4.98
Impervious	75	201	26	24	47	912	0	1285	70.97 ± 2.48
Bareland	9	0	6	2	0	0	17	34	50.00 ± 16.81
Total	4120	2377	474	349	341	1158	76	8895	
PA (%)	94.51 ± 0.70	67.23 ± 1.89	82.28 ± 3.44	28.08 ± 4.71	56.89 ± 5.26	78.76 ± 2.36	22.37 ± 9.37		
OA (%) = 79.85 ± 0.83									

Table A4. Pixel level agreements between the GlobeLand30-2020 product and field plot data in northeast China and their ranges at the 95% confidence interval.

	Cropland	Forest	Grassland	Wetland	Water	Impervious	Bareland	Total	UA (%)
Cropland	3712	431	13	112	67	201	24	4560	81.40 ± 1.13
Forest	80	1281	8	50	22	20	5	1466	87.38 ± 1.70
Grassland	55	418	397	59	51	35	17	1032	38.47 ± 2.97
Wetland	13	8	7	80	9	0	0	117	68.38 ± 8.43
Water	2	14	4	18	151	0	0	189	79.89 ± 5.71
Impervious	252	223	32	20	36	900	16	1479	60.85 ± 2.49
Bareland	6	2	13	10	5	2	14	52	26.92 ± 12.06
Total	4120	2377	474	349	341	1158	76	8895	
PA (%)	90.10 ± 0.91	53.89 ± 2.00	83.76 ± 3.32	22.92 ± 4.41	44.28 ± 5.27	77.72 ± 2.40	18.42 ± 8.72		
OA (%) = 73.47 ± 0.92									

Table A5. Pixel level agreements between the GLC_FCS30-2020 product and field plot data in northeast China and their ranges at the 95% confidence interval.

	Cropland	Forest	Grassland	Wetland	Water	Impervious	Bareland	Total	UA (%)
Cropland	3873	586	71	117	43	193	25	4908	78.91 ± 1.14
Forest	71	1391	10	61	15	16	4	1568	88.71 ± 1.57
Grassland	118	212	366	43	13	19	17	788	46.45 ± 3.48
Wetland	8	12	3	80	30	3	5	141	56.74 ± 8.18
Water	5	6	0	16	194	3	1	225	86.22 ± 4.50
Impervious	42	170	21	27	46	924	11	1241	74.46 ± 2.43
Bareland	3	0	3	5	0	0	13	24	54.17 ± 19.93
Total	4120	2377	474	349	341	1158	76	8895	
PA (%)	94.00 ± 0.73	58.52 ± 1.98	77.22 ± 3.78	22.92 ± 4.41	56.89 ± 5.26	79.79 ± 2.31	17.11 ± 8.47		
OA (%) = 76.91 ± 0.88									

Table A6. Pixel level agreements between the CLCD-2020 product and field plot data in northeast China and their ranges at the 95% confidence interval.

	Cropland	Forest	Grassland	Wetland	Water	Impervious	Bareland	Total	UA (%)
Cropland	3674	575	59	117	56	205	28	4714	77.94 ± 1.18
Forest	143	1304	133	83	29	109	8	1809	72.08 ± 2.07
Grassland	77	241	236	33	8	42	9	646	36.53 ± 3.71
Wetland	8	8	0	63	62	3	0	144	43.75 ± 8.10
Water	0	0	0	0	105	0	0	105	100.00 ± 0.00
Impervious	218	249	46	48	81	797	22	1461	54.55 ± 2.55
Bareland	0	0	0	5	0	2	9	16	56.25 ± 24.31
Total	4120	2377	474	349	341	1158	76	8895	
PA (%)	89.17 ± 0.95	54.86 ± 2.00	49.79 ± 4.50	18.05 ± 4.04	30.79 ± 4.90	68.83 ± 2.67	11.84 ± 7.26		
OA (%) = 69.57 ± 0.96									

Table A7. Pixel level agreements between the ESA10-2020 product and field plot data in northeast China and their ranges at the 95% confidence interval.

	Cropland	Forest	Grassland	Wetland	Water	Impervious	Bareland	Total	UA (%)
Cropland	3563	396	72	86	39	245	20	4421	80.59 ± 1.17
Forest	223	1396	20	43	22	71	9	1784	78.25 ± 1.91
Grassland	90	303	331	95	25	12	23	879	37.66 ± 3.20
Wetland	4	4	1	69	6	1	1	86	80.23 ± 8.42
Water	8	14	1	17	167	10	1	218	76.61 ± 5.62
Impervious	108	191	23	20	36	739	9	1126	65.63 ± 2.77
Bareland	124	73	26	19	46	80	13	381	3.41 ± 1.82
Total	4120	2377	474	349	341	1158	76	8895	
PA (%)	86.48 ± 1.04	58.73 ± 1.98	69.83 ± 4.13	19.77 ± 4.18	48.97 ± 5.31	63.82 ± 2.77	17.11 ± 8.47		
OA (%) = 70.58 ± 0.95									

Table A8. Pixel level agreements between the CAF-LC30 2020 product and field data collected over CTERN plots distributed across China and their ranges at the 95% confidence interval.

	Cropland	Forest	Grassland	Wetland	Water	Impervious	Bareland	Total	UA (%)
Cropland	4	108	6	20	52	0	1	191	2.09 ± 2.03
Forest	0	1999	0	10	9	0	0	2018	99.06 ± 0.42
Grassland	1	123	64	32	47	1	0	268	23.88 ± 5.10
Wetland	0	19	0	36	1	0	0	56	64.29 ± 12.55
Water	0	1	0	0	150	0	0	151	99.34 ± 1.29
Impervious	0	10	0	0	46	22	0	78	28.21 ± 9.99
Bareland	0	34	0	4	10	0	3	51	5.88 ± 6.46
Total	5	2294	70	102	315	23	4	2813	
PA (%)	80.00 ± 35.06	87.14 ± 1.37	91.43 ± 6.56	35.29 ± 9.27	47.62 ± 5.52	95.65 ± 8.34	75.00 ± 42.44		
OA (%) = 80.98 ± 0.82									

Table A9. Pixel level agreements between the GlobeLand30-2020 product and field data collected over CTERN plots distributed across China and their ranges at the 95% confidence interval.

	Cropland	Forest	Grassland	Wetland	Water	Impervious	Bareland	Total	UA (%)
Cropland	4	142	7	7	41	0	1	202	1.98 ± 1.92
Forest	0	1901	1	6	12	0	0	1920	99.01 ± 0.44
Grassland	1	199	62	47	68	0	1	378	16.40 ± 3.73
Wetland	0	14	0	36	11	0	0	61	59.02 ± 12.34
Water	0	2	0	0	130	0	0	132	98.48 ± 2.09
Impervious	0	11	0	0	47	23	0	81	28.40 ± 9.82
Bareland	0	25	0	6	6	0	2	39	5.13 ± 6.92
Total	5	2294	70	102	315	23	4	2813	
PA (%)	80.00 ± 35.06	82.87 ± 1.54	88.57 ± 7.45	35.29 ± 9.27	41.27 ± 5.44	100.00 ± 0.00	50.00 ± 49.00		
OA (%) = 76.72 ± 0.88									

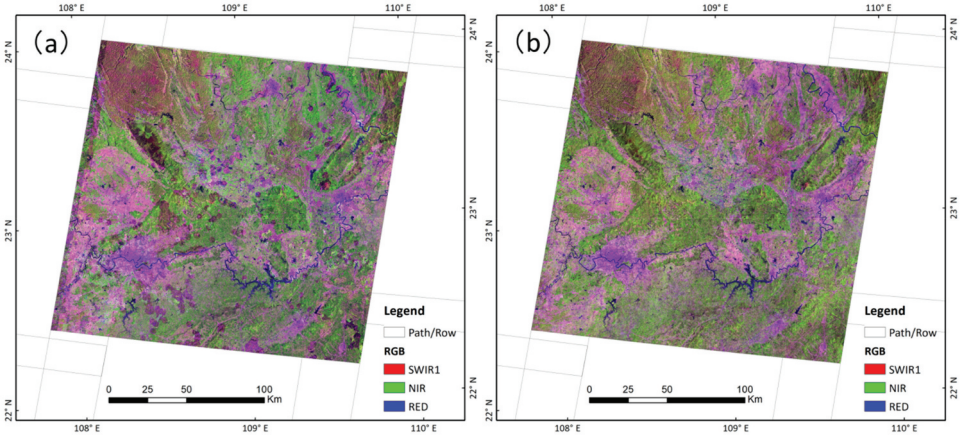

Figure A2. A clear view composite created using (a) the best available pixel (BAP) method and (b) the weighted BAP method used in this study. This area is covered by Landsat Worldwide Reference System (WRS-2) path 125 and row 44.

Table A10. Pixel level agreements between the GLC_FCS30-2020 product and field data collected over CTERN plots distributed across China and their ranges at the 95% confidence interval.

	Cropland	Forest	Grassland	Wetland	Water	Impervious	Bareland	Total	UA (%)
Cropland	3	144	6	16	90	0	0	259	1.16 ± 1.30
Forest	0	1930	0	10	9	0	2	1951	98.92 ± 0.46
Grassland	1	153	62	38	50	0	0	304	20.39 ± 4.53
Wetland	0	32	0	35	7	0	0	74	47.30 ± 11.38
Water	0	3	0	0	96	0	0	99	96.97 ± 3.38
Impervious	0	8	0	0	54	23	0	85	27.06 ± 9.44
Bareland	1	24	2	3	9	0	2	41	4.88 ± 6.59
Total	5	2294	70	102	315	23	4	2813	
PA (%)	60.00 ± 42.94	84.13 ± 1.50	88.57 ± 7.45	34.31 ± 9.21	30.48 ± 5.08	100.00 ± 0.00	50.00 ± 49.00		
OA (%) = 76.47 ± 0.88									

Table A11. Pixel level agreements between the CLCD-2020 product and field data collected over CTERN plots distributed across China and their ranges at the 95% confidence interval.

	Cropland	Forest	Grassland	Wetland	Water	Impervious	Bareland	Total	UA (%)
Cropland	3	126	6	12	84	2	1	234	1.28 ± 1.44
Forest	0	1957	3	11	15	0	0	1986	98.54 ± 0.53
Grassland	0	62	50	15	9	0	1	137	36.50 ± 8.06
Wetland	2	106	11	60	54	0	1	234	25.64 ± 5.59
Water	0	8	0	1	127	0	0	136	93.38 ± 4.18
Impervious	0	0	0	0	0	21	0	21	100.00 ± 0.00
Bareland	0	35	0	3	26	0	1	65	1.54 ± 2.99
Total	5	2294	70	102	315	23	4	2813	
PA (%)	60.00 ± 42.94	85.31 ± 1.45	71.43 ± 10.58	58.82 ± 9.55	40.32 ± 5.42	91.30 ± 11.52	25.00 ± 42.44		
OA (%) = 78.88 ± 0.85									

Table A12. Pixel level agreements between the ESA10-2020 product and field data collected over CTERN plots distributed across China and their ranges at the 95% confidence interval.

	Cropland	Forest	Grassland	Wetland	Water	Impervious	Bareland	Total	UA (%)
Cropland	4	156	5	18	66	1	1	251	1.59 ± 1.55
Forest	0	1901	1	14	13	0	0	1929	98.55 ± 0.53
Grassland	1	194	60	62	75	0	0	392	15.31 ± 3.56
Wetland	0	1	4	8	10	0	1	24	33.33 ± 18.86
Water	0	3	0	0	51	0	0	54	94.44 ± 6.11
Impervious	0	19	0	0	94	22	0	135	16.30 ± 6.23
Bareland	0	20	0	0	6	0	2	28	7.14 ± 9.54
Total	5	2294	70	102	315	23	4	2813	
PA (%)	80.00 ± 35.06	82.87 ± 1.54	85.71 ± 8.20	7.84 ± 5.22	16.19 ± 4.07	95.65 ± 8.34	50.00 ± 49.00		
OA (%) = 72.80 ± 0.92									

Table A13. Statistical significant of the fitting performance for 2020 land cover products with the 9th NFI data.

Product	Variable	Estimate	Standard Error	t Value	p Value
CAF-LC30 2020	Intercept (b0)	−9.271	1.750	−5.507	6.21e-06***
	Slope (b1)	1.189	0.045	26.686	<2e-16***
GlobeLand30-2020	Intercept (b0)	−8.823	2.010	−4.390	1.37e-04***
	Slope (b1)	1.175	0.052	22.81	<2e-16***
GLC_FCS30-2020	Intercept (b0)	−7.142	3.968	−1.800	0.0823*
	Slope (b1)	1.104	0.102	10.85	1.01e-11***
CLCD-2020	Intercept (b0)	−11.772	2.065	−5.700	3.63e-06***
	Slope (b1)	1.382	0.053	26.100	<2e-16***
ESA10-2020	Intercept (b0)	−12.605	3.218	−3.917	5.01e-04***
	Slope (b1)	1.517	0.083	18.392	<2e-16***

Significant level: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$