Science Bulletin 65 (2020) 1125–1136

Contents lists available at ScienceDirect

Science Bulletin

journal homepage: www.elsevier.com/locate/scib

Article

An updated Vegetation Map of China (1:1000000)

Yanjun Su [a,b], Qinghua Guo [a,b,*], Tianyu Hu [a,b], Hongcan Guan [a,b], Shichao Jin [a,b], Shazhou An [c], Xuelin Chen [d], Ke Guo [a,b], Zhanqing Hao [e], Yuanman Hu [e], Yongmei Huang [f], Mingxi Jiang [g], Jiaxiang Li [h], Zhenji Li [i], Xiankun Li [j], Xiaowei Li [k], Cunzhu Liang [l], Renlin Liu [m], Qing Liu [n,o], Hongwei Ni [p], Shaolin Peng [q], Zehao Shen [r], Zhiyao Tang [r], Xingjun Tian [s], Xihua Wang [t], Renqing Wang [u], Zongqiang Xie [a,b], Yingzhong Xie [k], Xiaoniu Xu [v], Xiaobo Yang [w], Yongchuan Yang [x,y], Lifei Yu [z], Ming Yue [aa], Feng Zhang [bb,cc], Keping Ma [a,b]

[a] State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China
[c] College of Pratacultural and Environmental Sciences, Xinjiang Agricultural University, Urumqi 830052, China
[d] College of Life Science, Northwest Normal University, Lanzhou 730070, China
[e] Key Laboratory of Forest Ecology and Management, Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110016, China
[f] College of Resources Science and Technology, Beijing Normal University, Beijing 100875, China
[g] Key Laboratory of Aquatic Botany and Watershed Ecology, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China
[h] College of Forestry, Central South University of Forestry & Technology, Changsha 410004, China
[i] College of Life Sciences, Xiamen University, Xiamen 361005, China
[j] Guangxi Institute of Botany, Chinese Academy of Sciences, Guilin 541006, China
[k] Agricultural College, Ningxia University, Yinchuan 750021, China
[l] School of Ecology and Environment, Inner Mongolia University, Hohhot 010000, China
[m] Institute of Navel Orange, School of Life and Environmental Science, Gannan Normal University, Ganzhou 341000, China
[n] Key Laboratory of Mountain Ecological Restoration and Bioresource Utilization, Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu 610041, China
[o] Ecological Restoration Biodiversity Conservation Key Laboratory of Sichuan Province, Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu 610041, China
[p] Institute of Natural Resources and Ecology, Heilongjiang Academy of Sciences, Harbin 150040, China
[q] State Key Laboratory of Biocontrol and Guangdong Key Laboratory of Plant Resources, School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China
[r] Institute of Ecology, College of Urban and Environmental Science and Key Laboratory of Earth Surface Processes, Peking University, Beijing 100871, China
[s] School of Life Sciences, Nanjing University, Nanjing 210023, China
[t] Zhejiang Tiantong Forest Ecosystem National Observation and Research Station, School of Ecological and Environmental Sciences, East China Normal University, Shanghai 200062, China
[u] School of Life Sciences, Shandong University, Qingdao 266237, China
[v] Department of Forest Science, Anhui Agricultural University, Hefei 230036, China
[w] Institute of Tropical Agriculture and Forestry, College of Material and Chemical Engineering, Hainan University, Haikou 570228, China
[x] Key Laboratory of Three Gorges Reservoir Region's Eco-Environment, Ministry of Education, Chongqing University, Chongqing 400045, China
[y] National Centre for International Research of Low-carbon and Green Buildings, Chongqing University, Chongqing 400045, China
[z] College of Life Science, Guizhou University, Guiyang 550025, China
[aa] Key Laboratory of Resource Biology and Biotechnology in Western China, Northwest University, Xi'an 710127, China
[bb] Institute of Loess Plateau, Shanxi University, Taiyuan 030006, China
[cc] School of Life Science, Shanxi University, Taiyuan 030006, China

ARTICLE INFO

Article history:
Received 14 November 2019
Received in revised form 12 February 2020
Accepted 18 February 2020
Available online 2 April 2020

Keywords:
Vegetation map
Crowdsource
Remote sensing
Update

ABSTRACT

Vegetation maps are important sources of information for biodiversity conservation, ecological studies, vegetation management and restoration, and national strategic decision making. The current Vegetation Map of China (1:1000000) was generated by a team of more than 250 scientists in an effort that lasted over 20 years starting in the 1980s. However, the vegetation distribution of China has experienced drastic changes during the rapid development of China in the last three decades, and it urgently needs to be updated to better represent the distribution of current vegetation types. Here, we describe the process of updating the Vegetation Map of China (1:1000000) generated in the 1980s using a "crowdsourcing-change detection-classification-expert knowledge" vegetation mapping strategy. A total of 203,024 field samples were collected, and 50 taxonomists were involved in the updating process. The resulting updated map has 12 vegetation type groups, 55 vegetation types/subtypes, and 866 vegetation formation/sub-formation types. The overall accuracy and kappa coefficient of the updated map are 64.8%

* Corresponding author.
E-mail address: qguo@ibcas.ac.cn (Q. Guo).

https://doi.org/10.1016/j.scib.2020.04.004

and 0.52 at the vegetation type group level, 61% and 0.55 at the vegetation type/subtype level and 40% and 0.38 at the vegetation formation/sub-formation level. When compared to the original map, the updated map showed that 3.3 million km^2 of vegetated areas of China have changed their vegetation type group during the past three decades due to anthropogenic activities and climatic change. We expect this updated map to benefit the understanding and management of China's terrestrial ecosystems.

1. Introduction

Vegetation maps provide crucial information for ecological studies, biodiversity conservation, vegetation management and restoration, and national strategic decision making [1–5]. China contains large diversity in vegetation types due to its vast topological and climatic variability [6,7], and accurately mapping the distribution of different vegetation types of China has been an essential but difficult task for the Chinese scientific community. The current Vegetation Map of China was produced by a team of more than 250 scientists over the course of two decades starting in the 1980s, who spent a huge amount of efforts to manually delineate vegetation formation/sub-formation boundaries and recognize vegetation formation/sub-formation types in the field [8]. This work fulfilled the needs of a national vegetation map of China at a scale of 1:1000000, and has been used as a fundamental input for policy making and ecological studies since its publication [9–12]. However, China has experienced rapid development during the last three decades [13–15], and its vegetation distribution has changed drastically due to climate change and anthropogenic activities (e.g., urban development, afforestation and management of desertification) [16–19]. The Vegetation Map of China urgently needs to be updated to improve the understanding and management of China's terrestrial ecosystems.

Traditional vegetation maps have mapped the potential vegetation distributions through the combination of field surveys, expert knowledge, and literature reviews [20–24]. With the development of remote sensing technologies, it has become possible to collect large-scale land surface observations in an efficient and repeatable way [25–27], which has greatly changed the vegetation mapping strategy [28,29]. Until recently, the most common method for vegetation mapping has been manually delineating units of vegetation formations and interpreting their attributes from very high-resolution aerial photos. For example, the United States of America, Japan, and certain European countries have adopted this method to generate their national vegetation maps [30]. However, collecting nationwide aerial photos is very expensive, especially for a country with an area as large as China. Recent advances in airborne hyperspectral and light detection and ranging (LiDAR) technologies provide new opportunities for vegetation mapping. The detailed spectral and structural information provided by the airborne hyperspectral and LiDAR data has the potential to be used to delineate vegetation formations and identify their types automatically and accurately [31,32], which can result in great manpower savings compared to the traditional manual drawing methods. Unfortunately, existing studies are still limited in scale, and complete national coverages of hyperspectral and LiDAR data are not yet available for most countries.

Spaceborne remote sensing improves the availability of datasets for large-scale vegetation mapping [33,34]. Numerous studies have been conducted to identify vegetation type attributes from satellite images using pixel-based or object-based classification schemes [35,36], and it has been found that object-based classification schemes perform better at delineating the boundaries between vegetation patches and identifying their types [37,38]. However, because of the similarity in the reflectance spectra among certain vegetation types, it is still difficult to map vegetation formations from spaceborne images, and most current studies only focus on mapping land-cover types (e.g., grassland, coniferous forest, deciduous forest, and crops) [39–42]. Recent studies have revealed that vegetation phenology information can be very helpful for separating vegetation patches and identify their formation attributes [43,44]. Moreover, with the accumulation of time-series images from different satellite platforms (e.g., Landsat, Sentinel, MODIS, and AVHRR) and the development of spatiotemporal data fusion techniques, it is now possible to generate unified time-series of high-resolution satellite images [45–47] and thereby accurately estimate vegetation phenology parameters [48,49]. This provides new opportunities for large-scale vegetation mapping. Nevertheless, to the best of our knowledge, none of the current studies have tried to update the Vegetation Map of China through the use of time-series satellite images. One of the major difficulties for this type of effort is the lack of field samples. The current Vegetation Map of China consists of over 800 formations and sub-formation groups, and a large amount of field samples are needed both for training a classifier and for validating the vegetation mapping results.

This study aimed to update the Vegetation Map of China (1:1000000) by adopting a "crowdsourcing-change detection-classification-expert knowledge" vegetation mapping strategy. A mobile phone application and an online web system were developed to collect and manage crowdsourced field samples. Together with field plots, a total of 203,024 field samples were collected. A change detection procedure between the current Vegetation Map of China and the land-cover map was used to detect areas with changed vegetation types, and a machine learning-based classification procedure was used to determine their new vegetation types. Finally, 50 taxonomists were involved in the validation and correction of the updated vegetation map based on their expert knowledge. We believe that the updated Vegetation Map of China can better reflect the current vegetation distribution status of China, and aid vegetation management, vegetation restoration and biodiversity conservation.

2. Materials and methods

2.1. Field sample data

Field samples are the foundation of vegetation mapping. In this study, three types of field samples were collected, including crowdsourced samples, plot samples, and literature samples (Fig. 1). The details of each type of field samples are described below.

2.1.1. Crowdsourced samples

We developed a mobile phone application (supporting both Android and iOS operating systems) to assist with the collection of crowdsourced samples (Fig. S1 online). The application has functions for viewing and downloading satellite images in areas of interests, uploading vegetation photos taken by the phone, recording vegetation formation or sub-formation attributes, viewing and

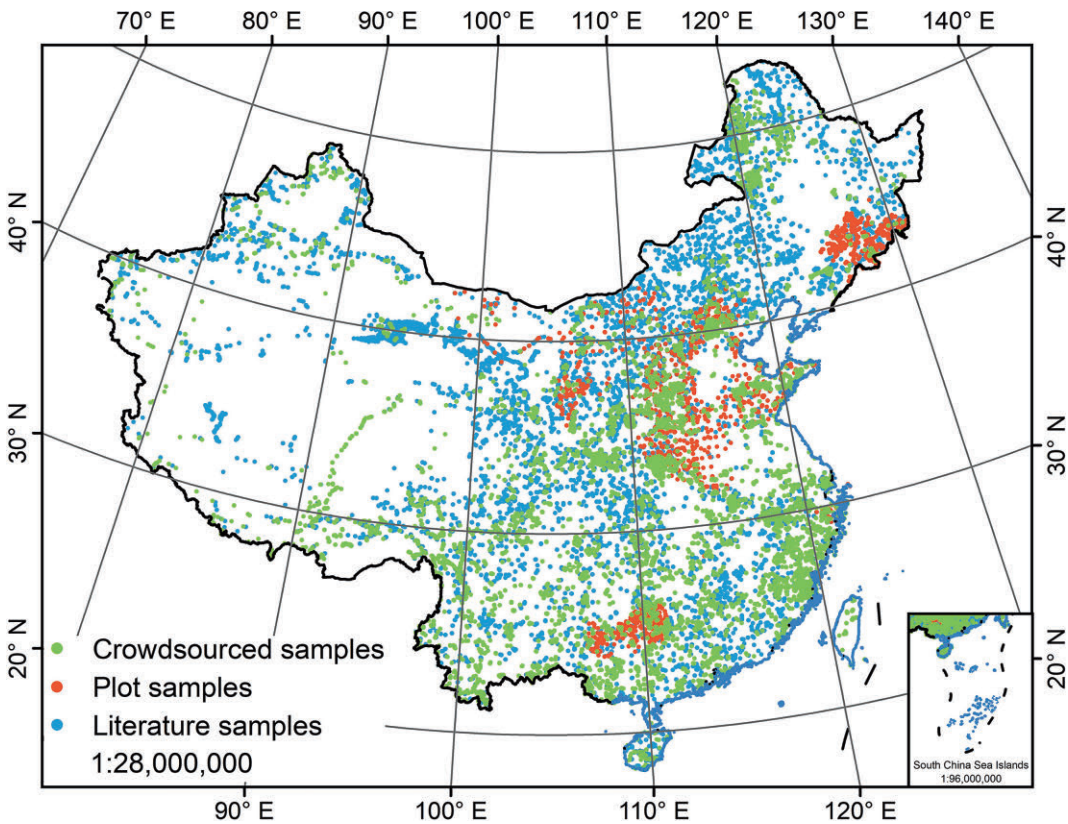

Fig. 1. Distribution of collected field samples used for updating the Vegetation Map of China.

managing vegetation formation/sub-formation segments, planning and tracking sampling routes in the field, adding points of interest, navigation based on planned sampling routes or points of interest, and so on. An online web system was also developed to manage the crowdsourced samples collected from the field, provide protocols for taxonomists to validate the collected photos, upload field samples from other sources, sync points of interest, sampling routes and vegetation segments with the mobile phone application, and so on (Fig. S1 online). At the time of this writing, there are over 1200 registered users, and over 177,580 crowdsourced samples collected (Fig. 1). The vegetation formation/sub-formation attributes of these collected crowdsourced samples were examined by 50 taxonomists and used in the following vegetation mapping process.

2.1.2. Plot samples

Field vegetation surveys were performed from 2010 to 2018, and overall 4711 field plots have been investigated (Fig. 1). Each field plot was set up following three rules: (1) it should be in a natural vegetation site that is rarely disturbed; (2) it should have relatively constant vegetation species composition; and (3) it should be representative of the corresponding vegetation formation/sub-formation segment. During the field vegetation surveys, the vegetation species composition and the location of each plot were recorded, and taxonomists determined the corresponding vegetation formation/sub-formation attributes.

2.1.3. Literature samples

In addition to the above field samples, we also collected 17,749 samples from the literature (Fig. 1). The reviewed literature included scientific papers and biodiversity study books of natural reserves published in the last 10 years. Records with specific species composition information and geolocation information were retained and digitized, and taxonomists then determined their formation/sub-formation attributes.

2.2. Remote sensing datasets

In this study, we used data from four remote sensing datasets for updating the Vegetation Map of China (1:1000000) (Table S1 online). MOD13A2 is a MODIS product providing vegetation indices (including the normalized difference vegetation index/NDVI and the enhanced vegetation index/EVI) at a spatial resolution of 1 km every 16 days [50]. The maximum value composite method based on NDVI is used to generate the product from daily MODIS images during the 16-day composite period. Here we used all EVI data for the year 2015 from the MOD13A2 product Version 6. MCD12Q2 is also a MODIS product and provides global land surface phenology metrics at yearly intervals [51]. In this study, the MOD12Q2 product Version 6 of the year 2015 was used. All vegetation phenological variables included in the product, that are NumCycles, Greenup, MidGreenup, Peak, Maturity, MidGreendown, Senescence, Dormancy, EVI_Minimum, EVI_Amplitude, EVI_Area, were used in the vegetation mapping process. The definition of each phenological variable is provided in Table S1 (online).

Four terrain products, i.e., elevation, slope, aspect, and roughness, were derived from the corrected Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) product provided by Zhao et al. [52]. The SRTM DEM is one of most widely used global DEM datasets [53,54]. However, it was found that the original SRTM DEM systematically overestimates terrain elevation in vegetated areas [55,56]. Zhao et al. [52] generated a global corrected SRTM DEM product through the fusion of airborne LiDAR data, spaceborne LiDAR data and optical imagery, which effectively reduced the systematic bias of the original SRTM DEM product.

The terrain slope, aspect and roughness data were generated from the corrected SRTM DEM data using the Esri ArcGIS software (version 10.3; https://www.esri.com/). We also used a land cover of China for 2010 extracted from the global land cover dataset by Chen et al. [46]. This map was generated integrating pixel- and object-based methods with knowledge at 30 m resolution, and includes 10 land cover types (i.e., cultivated land, forest, grassland, shrubland, water bodies, wetland, tundra, artificial surfaces, bareland and permanent snow and ice).

2.3. Ancillary datasets

In addition to the above-mentioned remote sensing datasets, two climate products, i.e., the climate zones of China and climate surfaces, were also used in this study. The climate zones of China were created by China's Meteorological Administration in 1978. It divides China into 10 climate zones and 32 sub-zones (Fig. S2 online) based on the climate data recorded from 1951 to 1970. Moreover, 24 climate surfaces, including average monthly mean temperature and average monthly precipitation from January to December, were calculated at a resolution of 1 km from the monthly WorldClim datasets from 1970 to 2000. In this study, we used the WorldClim Version 2 product, which was interpolated from weather station data using the thin-plate splines with covariates including elevation, distance to the coast and satellite-derived variables [57].

2.4. Detect changed areas in vegetation type group

As mentioned, China has experienced rapid development during the past three decades, and its vegetation distribution has changed drastically during the development [16–19]. In this study, we made the assumption that a vegetation formation/sub-formation type did not change in an area if the vegetation type group had not changed. Therefore, areas with unchanged vegetation type groups were directly retained in the updated Vegetation Map of China, and determining areas with changed vegetation type groups was the essential first step of the vegetation mapping process.

The Vegetation Map of China has three classification levels: (1) vegetation type group (Table S2 online), (2) vegetation type/subtype (Table S3 online), and (3) vegetation formation/sub-formation (Table S4 online). In this study, we crosswalked the vegetation type groups to the 30 m land cover classes (Table S2 online) and then compared the original Vegetation Map to the land cover of China. To perform the crosswalked comparison, we first converted the original Vegetation Map of China to a raster file with the same spatial resolution (i.e., 30 m) of the land cover map. If the vegetation type group of a pixel could be matched to the corresponding land cover type, we assumed that the corresponding formation/sub-formation type did not change as well, and the pixel was treated as unchanged. Finally, the raster file of changed/unchanged areas was converted back to a polygon file using the conversion tool in ESRI ArcGIS software. All changed and unchanged areas were then sorted by overlaying them with the climate zone distribution map.

2.5. Update Vegetation Map of China

A random forest-based procedure was used to update the Vegetation Map of China in each climate zone, separately (Fig. 2). Within a climate zone, the unchanged areas derived from the change detection step were first grouped based on their land cover types. Then, a random forest classifier was built separately for every land cover type in the climate zone by using the vegetation formation/sub-formation types of the unchanged areas with the same land cover type as training samples. All random forest classifiers in each climate zone were combined into vegetation mapping models. During the prediction process of a changed area in a climate zone, a classifier was first selected from the model by matching the land cover type, and this matched classifier was used to predict the vegetation formation/sub-formation type. In this study, the random forest method was implemented using the RandomForestClassifier function of the Scikit-learn package in Python [58]. All remote sensing variables listed in Table S1 (online) and climate surfaces were used as predictors in the random forest method. The update process was operated based on vegetation segments, and the average values of the predictors in each segment were fed into the random forest algorithm. The two parameters in the random forest algorithm, i.e., *n*tree (number of trees) and *m*try (number of variables tried at each split), were set as 500 and 4, respectively. Note that if the land cover type of an area changed into water bodies, artificial surfaces, or permanent snow and ice, it was directly updated as non-vegetated area, skipping the random forest classification process described above.

The unchanged areas and modified changed areas were merged together to generate a new pre-calibrated Vegetation Map of China. This map was sent to 50 taxonomists very familiar with the vegetation distribution in their local areas for validation and calibration. They visually examined the pre-calibrated map and looked for any potential errors in the map based on their knowledge. If any errors were found in the pre-calibrated map, they were manually corrected based on expert knowledge or field vegetation survey using the mobile phone application and online web system developed for this study. Among the 177,580 collected crowdsourced samples, around 160,000 samples were collected and used in this calibration process. Moreover, the 17,749 field samples from the literature were also used. The final calibrated map was used as the updated Vegetation Map of China (1:1000000).

2.6. Accuracy assessment

The accuracy of the updated Vegetation Map of China was evaluated at the vegetation type/subtype level and the vegetation formation/sub-formation level, respectively. The remaining ~ 25,000 field samples that were not used in the calibration procedure were used as reference. To ensure the representativeness of the samples used for validation, we further filtered them to ensure that there were at least 500 samples for each vegetation type/subtypes. Moreover, vegetation types related to crop may change frequently in China due to crop rotation. Therefore, we further removed samples falling in cultivated lands of the land cover map to reduce the influence recent changes in vegetation type between 2014 and 2019. Finally, 10,766 field samples covering 10 vegetation type/subtypes were used to calculate the overall accuracy and kappa coefficient. If the species of a field sample was contained in the vegetation formation/sub-formation type of the updated Vegetation Map of China, it was counted as a match; otherwise, it was counted as a mismatch. We further divided these validation samples into two groups, falling either in changed areas or in unchanged areas. The overall accuracy and kappa coefficient were also calculated for each group to compare the accuracy of the updated and original Vegetation Maps of China.

3. Results

3.1. Changes in vegetation type group during the past three decades

By comparing the original Vegetation Map of China to the 30 m land cover map, we found that 3.3 million km^2 of vegetated areas changed their vegetation type group, accounting for around one third of the national land area of China. Southern China experi-

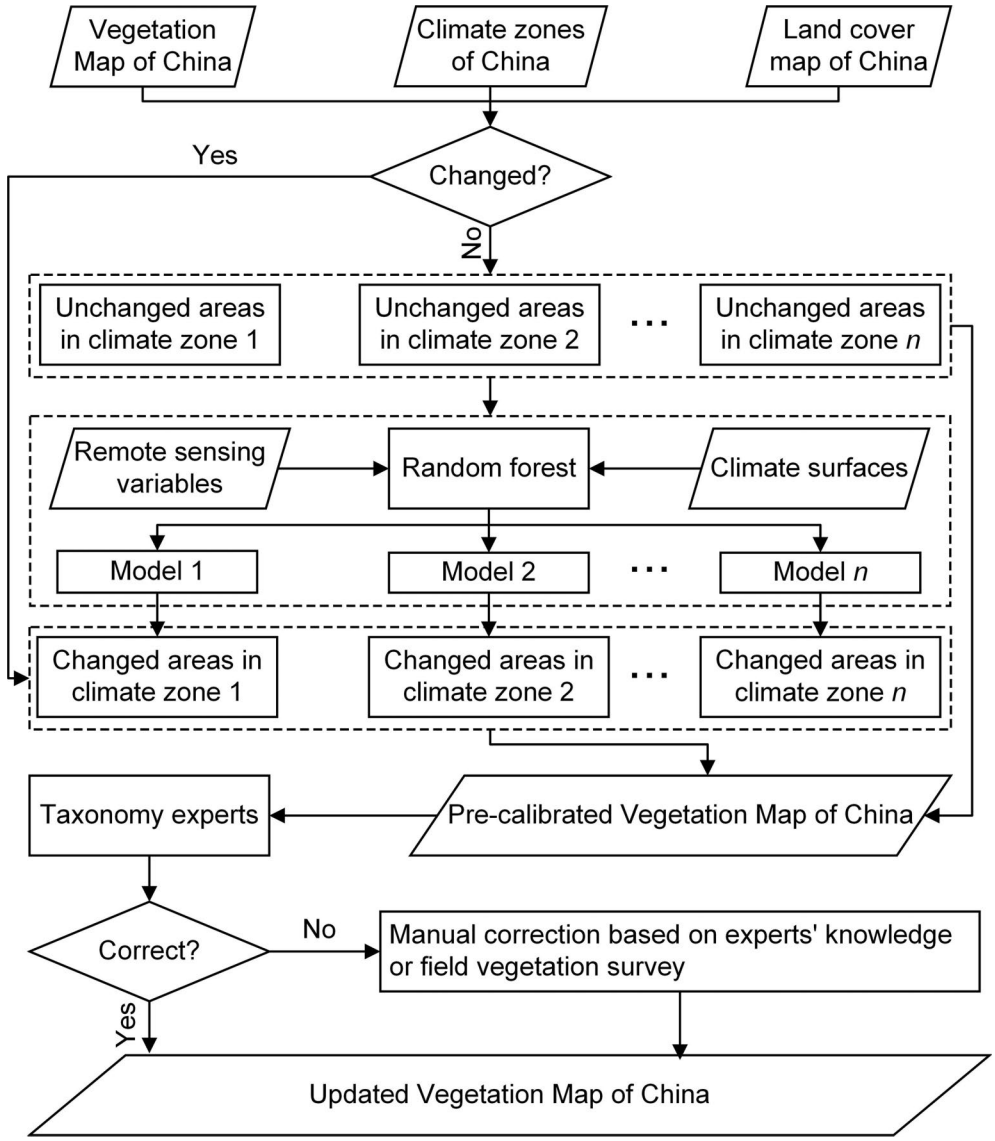

Fig. 2. The flow chart for updating the Vegetation Map of China.

enced the most severe vegetation change, while the vegetation in western and northern China was relatively better preserved (Fig. 3a). Among the 12 vegetation type groups, the changed areas of scrub and cultural vegetation were the largest, around 0.8 million km^2 each (Fig. 3b). Steppe, meadow, grass-forb community, broadleaf forest and needleleaf forest had a changed area of around 0.2–0.4 million km^2, and desert, swamp, alpine vegetation and mixed needleleaf and broadleaf forest had a changed area lower than 0.1 million km^2 (Fig. 3b). Over 95% of scrub in the original Vegetation Map of China changed vegetation type group (Fig. 3c). Moreover, although swamp and grass-forb community underwent a relatively small area change, their changed areas accounted for over 80% of their total areas (Fig. 3c).

Cultivated land, forests, grassland, non-vegetated area, and shrubland were the five land cover types that most vegetation type groups changed to (Fig. 4). Nearly all vegetation type groups, except alpine vegetation, had a large portion (>15%) changed to cultural vegetation. Scrub, swamp, meadow, grass-forb community, alpine vegetation, non-vegetated area, and cultural vegetation also had a portion larger than 10% changed to forest. Scrub, swamp cultural land, broadleaf forest, needleleaf forest, mixed needleleaf and broadleaf forest, and non-vegetated area had a portion larger than 25% changed to grassland. Steppe, alpine vegetation, meadow, desert, and cultural vegetation had a portion larger than 25% changed to non-vegetated area. Desert and alpine vegetation were the only two groups having more than 15% of the area changed to shrubland.

3.2. The updated Vegetation Map of China (1:1000000)

The updated Vegetation Map of China (1:1000000) at the vegetation type group level is shown in Fig. 5a. Northeastern China is dominated by needleleaf forest, cultural vegetation, and steppe; the North China Plain is mainly dominated by cultural vegetation; northwestern China is mainly covered by desert; the Tibetan Plateau is dominated by steppe, meadow, grass-forb community and alpine vegetation; southern China is much more fragmented, with broadleaf forest, cultural vegetation, and scrub. Compared to the original Vegetation Map of China, the area of broadleaf forest increases significantly, the areas of scrub and cultural vegetation decrease slightly, and the area of other vegetation groups remain

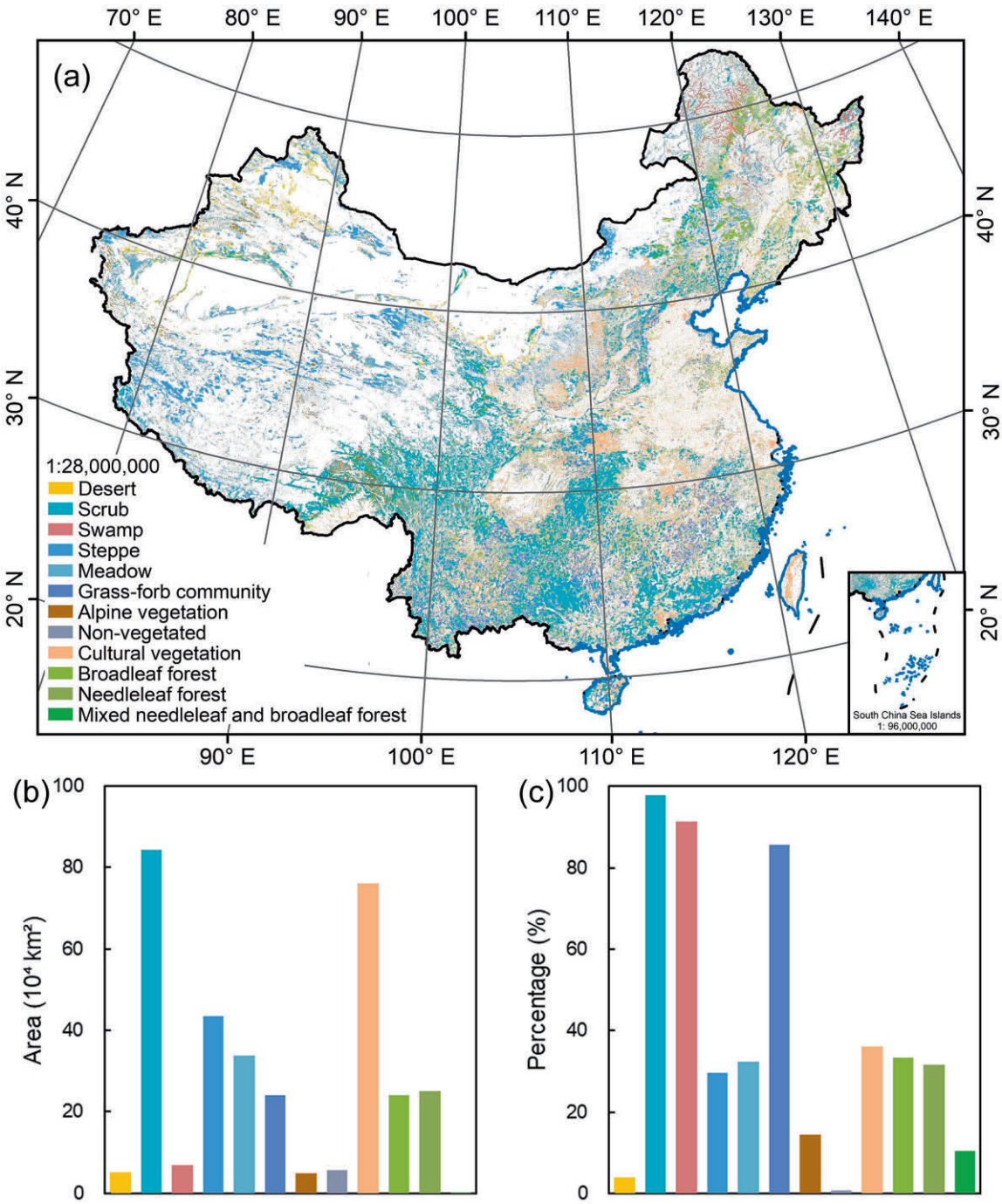

Fig. 3. Map of changed vegetation (a), changed area by class (b), and changed percentage relative to the original distribution of the Vegetation Map of China at the vegetation type group level by comparison with the land cover map of 2010 (c). The colors in subplot (b) and (c) correspond to the legend in subplot (a).

stable (Table S5 online). The non-vegetated area increases by 11,000 km^2 (Table S5 online).

The updated Vegetation Map of China (1:1000000) at the vegetation type/subtype level is shown in Fig. 5b. In northeastern China, the major vegetation type/subtype in the needleleaf forest group is the cold-temperate and temperate mountains needleleaf forest, the major vegetation type/subtype in the cultural vegetation group is the annual short growing period cold-resistant crops (without fruit trees), and the major vegetation type/subtype in the meadow group is the temperate needlegrass arid steppe. In North China Plain, the major vegetation type/subtype in the cultural vegetation group is the annually and cold-resistant economic crops. In northwestern China, the major vegetation type/subtype of the desert group is the temperate dwarf semi-arboreous desert. In the Tibetan Plateau, the major vegetation type/subtype in the steppe group is the Alpine grass, *Carex* steppe, and the major vegetation type/subtype in the alpine vegetation group is the Alpine *Kobresia* spp., forb meadow. In southern China, the vegetation formation/sub-formation composition is much more complicated, and is represented by a mixture of subtropical needleleaf forest, subtropical and tropical broadleaf evergreen and deciduous scrub (always with scattered trees), annual rotations of two or three crops containing upland and irrigation rotate crops annually (with double-cropping rice), evergreen orchards and subtropical economic forest, among others.

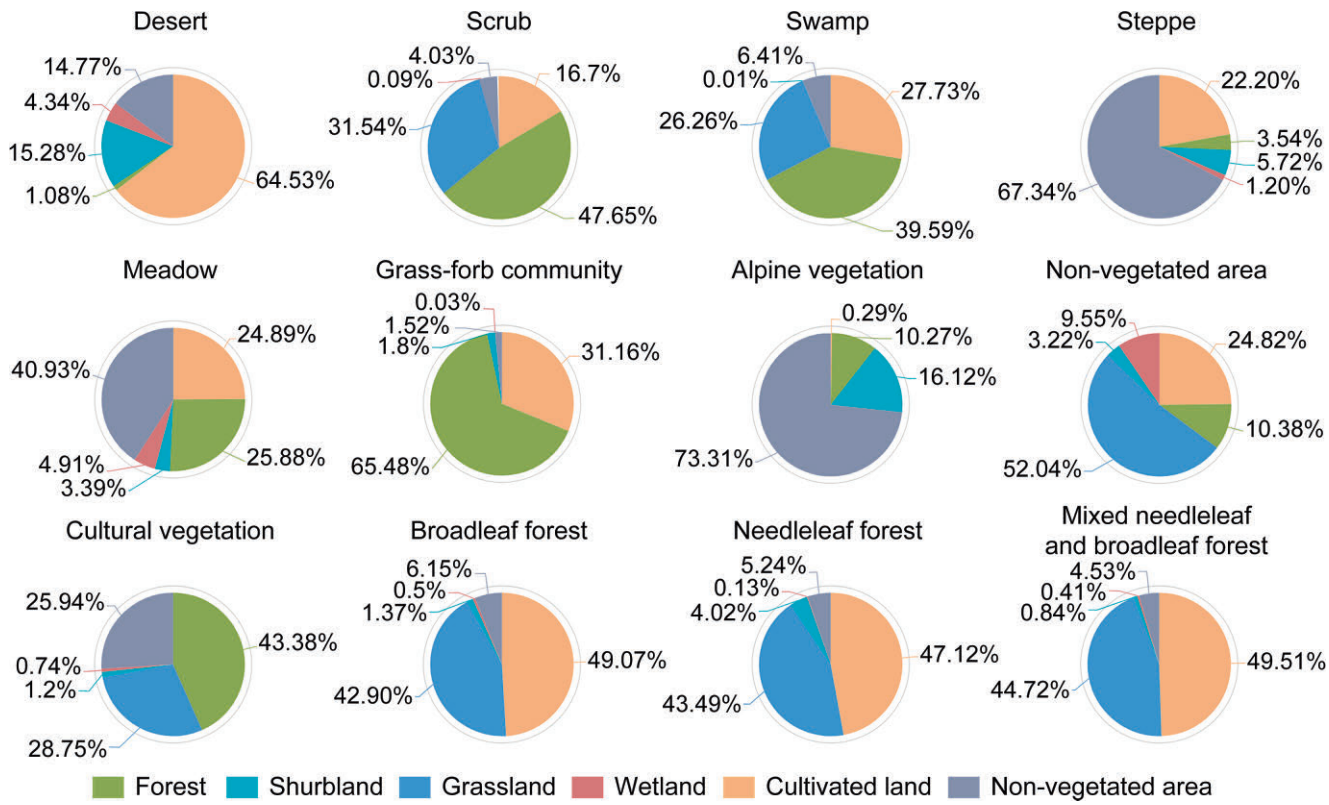

Fig. 4. Statistics of changes for each vegetation type group in the original Vegetation Map of China by comparing with the land cover map.

Due to space limitations, we selected six 300 km × 300 km examples across China to present the updated Vegetation Map of China at the formation/sub-formation level (Fig. 6). Generally, northern China (e.g., example areas 1 and 2) had larger vegetation formation/sub-formation patches, and the vegetation formation/sub-formation composition was simpler than in southern China (Fig. S3 online). The average patch sizes in the example areas 3 and 5 were much smaller than in examples areas 1 and 2.

The quantitative accuracy assessment results are shown in Table 1. At the vegetation type group level, the overall accuracy and kappa coefficient of the updated Vegetation Map of China were 64.8% and 0.52, respectively, comparing to values of 56.0% and 0.43 of the original Vegetation Map of China. The overall accuracy and kappa coefficient of the updated map in changed areas were improved by around 20% comparing to the original map. At the vegetation type/subtype level, the overall accuracy and kappa coefficient of the updated Vegetation Map of China were 60.8% and 0.55, respectively, while those of the original Vegetation Map of China were 50.5% and 0.44. The overall accuracy and kappa coefficient of the changed areas were 51.1% and 0.43, and were lower than those of unchanged areas (overall accuracy = 67.8%, kappa coefficient = 0.63). At the vegetation formation/sub-formation level, the overall accuracy and kappa coefficient of the updated Vegetation Map of China were much lower, 39.7% and 0.38, respectively, but those values for changed areas of the updated map were much higher than in the original map (Table 1).

4. Discussion

Comparing with the land cover map product generated by Chen et al. [46], the vegetation type group of nearly 35% of the national land area of China changed during this time period. Although some of the changes might be caused by errors in the original Vegetation Map of China and land cover map and mismatches between them, it still indicates that China has experienced many changes in vegetation composition during the rapid development of the last thirty years. Southern China, where the gross domestic product increased the fastest during the last thirty years [15], witnessed the largest changes (Fig. 3a). On the one hand, a lot of areas that were previously covered by natural vegetation have been converted to cropland; on the other hand, a lot of non-vegetated areas have been reforested. Along the coastline of southern China, a big portion of vegetated areas have been developed into urban areas. Northern and western China, which are less developed economically, underwent fewer changes in vegetation type group (Fig. 3a). In Inner Mongolia, most losses of grassland-related vegetation groups (e.g., steppe, grass-forb community, and meadow) may have been caused by grazing [56,59]. Meanwhile, the recent huge investments in afforestation and desertification management projects by the Chinese government have brought significant increases in the area of vegetation type groups related to forests and scrub and decreases in that of desert (Fig. 4) [60]. The vegetation of the Tibetan Plateau lost areas of grass-forb community, meadow and alpine vegetation, which might be the result of an integrated effect of climate changes and human activities in the region [61,62]. The North China Plain has long been one of the main crop-producing areas of China, and the vegetation composition in this region has been relatively stable (Fig. 3a).

To the best of our knowledge, this is the first study to produce and update the Vegetation Map of China using a "crowdsourcing-change detection-classification-expert knowledge" strategy. The mobile phone application and online web system developed for this study have been widely used as a field survey tool for vegetation mapping, which greatly improved the field vegetation survey efficiency. Unfortunately, although we collected over 170,000 crowdsourced field samples, they were still insufficient to cover all the vegetation formation/sub-formation types used in the

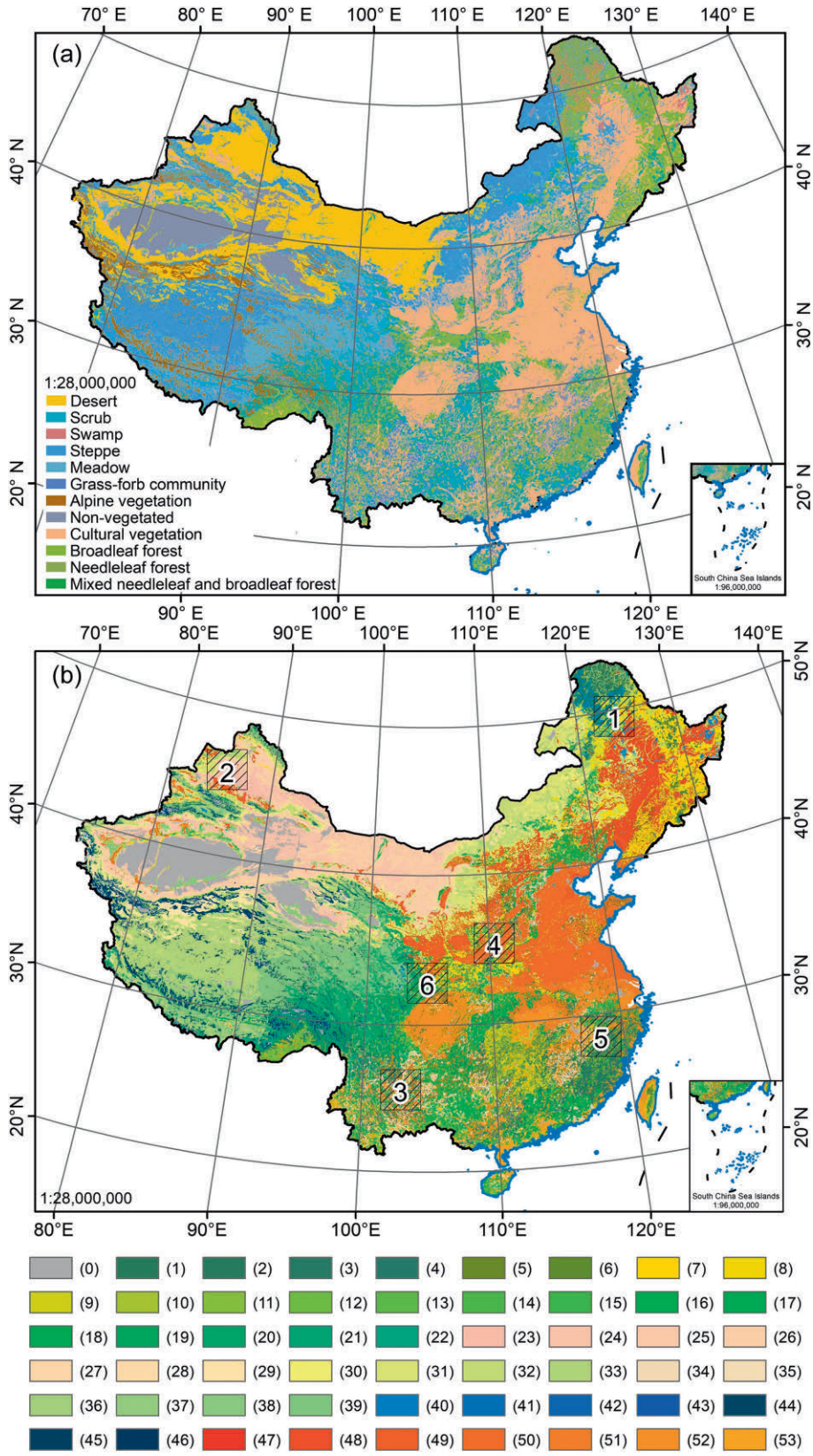

Fig. 5. The updated Vegetation Map of China (1:1000000) presented at the vegetation group type level (a) and at the vegetation type/subtype level (b). The six squares numbered from 1 to 6 in (b) represent six 300 km × 300 km examples used for presenting the updated Vegetation Map of China (1:1000000) at the formation/sub-formation level. The corresponding vegetation type/subtype name of each ID is listed in Table S3 (online).

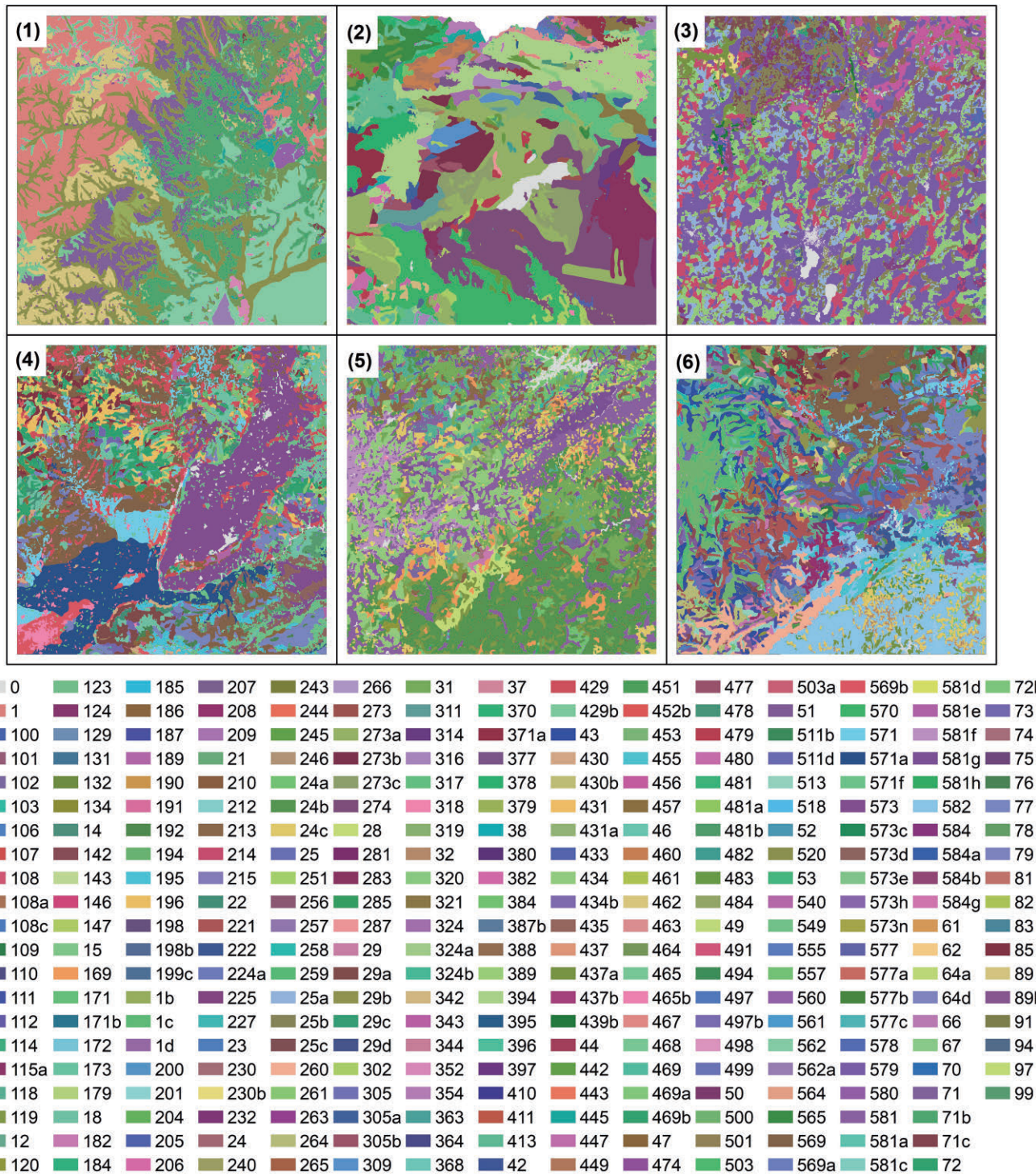

Fig. 6. Six examples of the updated Vegetation Map of China (1:1000000) presented at the vegetation formation/sub-formation level corresponding to the six square areas in Fig. 5b. The corresponding vegetation formation/sub-formation name of each ID is listed in Table S4 (online).

Vegetation Map of China. Therefore, we did not use field samples as training samples in the vegetation mapping process. Instead, we used a change detection method to first find the unchanged areas at the level of vegetation type group by comparing the original Vegetation Map of China with the land cover map from Chen et al. [46], and used the unchanged areas to build random forest-based models to predict the changed areas. However, there are a considerable amount of errors in the unchanged areas of the original Vegetation Map of China (Table 1), and vegetation indices and phenological parameters obtained from satellite images might be very similar among different vegetation formation/sub-formation types [39,40]. The resulting pre-calibrated vegetation map might still contain a large amount of errors, and we believe it is a still necessary to manually validate and correct the vegetation map. In this study, 50 taxonomists spent five months to manually validate and correct the updated map using over 160,000 field samples, which has greatly reduced the vegetation mapping errors from temporal discrepancies of different datasets, random forest models and so

Table 1
The accuracy of the original and updated Vegetation Map of China (1:1000000) at the vegetation type/subtype level and the vegetation formation/sub-formation level.

		Vegetation type group level		Vegetation type/subtype level		Vegetation formation/sub-formation level	
		Overall accuracy (%)	Kappa coefficient	Overall accuracy (%)	Kappa coefficient	Overall accuracy (%)	Kappa coefficient
Changed areas	Updated map	54.1	0.38	51.1	0.43	40.3	0.38
	Original map	33.1	0.20	26.8	0.21	5.0	0.04
Unchanged area	Updated map	72.6	0.61	67.8	0.63	39.3	0.38
	Original map	72.6	0.61	67.8	0.63	39.3	0.38
Overall	Updated map	64.8	0.52	60.8	0.55	39.7	0.38
	Original map	56.0	0.43	50.5	0.44	24.9	0.23

on. With the accumulation of field crowdsourced samples, we believe that the workload of the manual examinations could be greatly reduced in the future by employing deep learning-based classification methods [63–65].

The accuracy of the updated Vegetation Map of China has been much improved compared to the original map (Table 1). Since this study directly used the original vegetation map attributes in the unchanged areas, the accuracy of the updated map in the unchanged areas remains the same as in the original map. However, in the changed areas, the overall accuracy and kappa coefficient of the updated map are improved by over 20% and 18% at all three levels (Table 1), indicating that the original map needs to be updated urgently and the updated map may provide more up-to-date vegetation distribution information. Moreover, the accuracy of the updated map in the changed areas is very close to that of the original map in the unchanged areas, especially at the vegetation formation/sub-formation, indicating that the vegetation mapping strategy used in this study can generate accuracies equivalent to those of the original map.

Nevertheless, there are still several limitations in this study. First, a more complete accuracy assessment effort needs to be conducted. Due to the limited number of samples collected, the current study only used 10,766 samples covering 10 vegetation types/subtypes. Second, the classification system of the current Vegetation Map needs to be updated. The current classification system is mixed with vegetation formation and sub-formation information, and certain new vegetation formation types are not included [66]. Third, the accuracy of the updated Vegetation Map of China is still relatively low. Even in the unchanged areas, the overall accuracy for the original vegetation map is only 39% at the vegetation formation/sub-formation level. Since this study used the unchanged areas as the training samples to build the classification models, the large amount of errors in the original map may bring considerable noise in the updated map. Nevertheless, we believe that the updated map is still an important vegetation map product since the accuracy of the update map is much higher than that of the original map, especially in the changed areas. Currently, the author team is working on making the new generation of Vegetation Map of China (1:500000). A whole new vegetation classification system based on the large amount of field samples will be proposed, and a new vegetation mapping strategy purely based on field samples, high-resolution satellite images, deep learning methods and expert knowledge will be used. With the accumulation of crowdsourced field samples, we believe that this strategy can ultimately result in a much more detailed and accurate vegetation map of China. However, it may still need at least five to six years before the new Vegetation Map of China being produced. Before then, we believe that the resulting updated map in this study could better represent the current vegetation status of China and therefore serve better in ecological studies and managements and national strategic decision making.

5. Conclusions

This study updated the Vegetation Map of China (1:1000000) through the use of a "crowdsourcing-change detection-classification-expert knowledge" strategy. A mobile phone application and an online web system were developed to assist with the collection of crowdsourced field samples. Overall, over 200,000 field samples were collected to validate and evaluate the vegetation mapping results. Moreover, over 50 taxonomists were involved in the vegetation mapping process to manually examine and correct the updated vegetation map. The resulting updated Vegetation Map of China (1:1000000) includes 12 vegetation type groups, 55 vegetation types/subtypes, and 866 vegetation formation/sub-formation types. Its overall accuracy and kappa coefficient are 64.8% and 0.52 at the vegetation type group level, 60.8% and 0.55 at the vegetation type/subtype level and 39.7% and 0.38 at the vegetation formation/sub-formation level, which are about 10%–15% higher than those of the original Vegetation Map of China. With the provided much up-to-date vegetation distribution information, we believe the resulting map has great potential to be used as a replacement vegetation map product for ecological studies and managements and national strategic decision making.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (XDA19050401). Maps in this article were reviewed by Ministry of Natural Resources of the People's Republic of China (GS(2020)1044)). The updated Vegetation Map of China will be available through the project website (http://www.chinavegetation.cn/#/).

Author contributions

Qinghua Guo and Keping Ma design the study; Yanjun Su, Qinghua Guo, Tianyu Hu, Hongcan Guan, and Shichao Jin developed the mobile phone application and online web system, and conducted the study; Ke Guo provided the literature sample data; Ke Guo, Zhanqing Hao, Yuanman Hu, Yongmei Huang, Mingxi Jiang, Jiaxiang Li, Zhenji Li, Xiankun Li, Xiaowei Li, Cunzhu Liang, Renlin Liu, Qing Liu, Hongwei Ni, Shaolin Peng, Zehao Shen, Zhiyao Tang, Xingjun Tian, Xihua Wang, Renqing Wang, Zongqiang Xie, Yingzhong Xie, Xiaoniu Xu, Xiaobo Yang, Yongchuan Yang, Lifei Yu, Ming Yue, and Feng Zhang validated the crowdsourced samples and validated the updated Vegetation Map of China (1:1000000); Yanjun Su, Qinghua Guo, Tianyu Hu, Hongcan Guan, and Shichao

Jin wrote the manuscript; and all authors involved in the revision of the manuscript.

Appendix A. Supplementary materials

Supplementary materials to this article can be found online at https://doi.org/10.1016/j.scib.2020.04.004.

References

[1] Gaston KJ, Blackburn TM, Goldewijk KK. Habitat conversion and global avian biodiversity loss. Proc R Soc B-Biol Sci 2003;270:1293–300.
[2] Guo Y, Schöb C, Ma W, et al. Increasing water availability and facilitation weaken biodiversity-biomass relationships in shrublands. Ecology 2019;100: e02624.
[3] Lai L, Huang X, Yang H, et al. Carbon emissions from land-use change and management in China between 1990 and 2010. Sci Adv 2016;2:e1601063.
[4] Lenihan JM, Drapek R, Bachelet D, et al. Climate change effects on vegetation distribution, carbon, and fire in California. Ecol Appl 2003;13:1667–81.
[5] Tang X, Zhao X, Bai Y, et al. Carbon pools in China's terrestrial ecosystems: new estimates based on an intensive field survey. Proc Natl Acad Sci USA 2018;115:4021–6.
[6] Liu Y, Duan M, Yu Z. Agricultural landscapes and biodiversity in China. Agric Ecosyst Environ 2013;166:46–54.
[7] Yu X. Biodiversity conservation in China: barriers and future actions. Int J Environ Stud 2010;67:117–26.
[8] Zhang X, Sun S, Yong S, et al. Vegetation map of the People's Republic of China (1:1000000). Beijing: Geology Publishing House; 2007.
[9] Cheng T, Zhang J, Zheng X, et al. Land cover's refined classification based on multi source of remote sensing information fusion: a case study of national geographic conditions census in China. Proc SPIE 2018;10611:1061103.
[10] Fang J, Chen A, Peng C, et al. Changes in forest biomass carbon storage in China between 1949 and 1998. Science 2001;292:2320–2.
[11] Wan JZ, Wang CJ, Qu H, et al. Vulnerability of forest vegetation to anthropogenic climate change in China. Sci Total Environ 2018;621:1633–41.
[12] Zhao D, Wu S, Yin Y, et al. Vegetation distribution on Tibetan Plateau under climate change scenario. Reg Environ Chang 2011;11:905–15.
[13] Allen F, Qian J, Qian M. Law, finance, and economic growth in China. J Financ Econ 2005;77:57–116.
[14] Chow GC. Capital formation and economic growth in China. Quart J Econ 1993;108:809–42.
[15] National Bureau of Statistics of China. Statistical communiqué of the People's Republic of China on the 2014 national economic and social development. 2015. Available online: http://www.stats.gov.cn/tjsj/zxfb/201502/t20150226_685799.html (accessed on 27 March 2020)..
[16] Gao Q, Guo Y, Xu H, et al. Climate change and its impacts on vegetation distribution and net primary productivity of the alpine ecosystem in the Qinghai-Tibetan Plateau. Sci Total Environ 2016;554–555:34–41.
[17] Peng S, Chen A, Xu L, et al. Recent change of vegetation growth trend in China. Environ Res Lett 2011;6:044027.
[18] Wang J, Wang K, Zhang M, et al. Impacts of climate change and human activities on vegetation cover in hilly southern China. Ecol Eng 2015;81:451–61.
[19] Zhang Y, Zhang C, Wang Z, et al. Vegetation dynamics and its driving forces from climate change and human activities in the three-river source region, China from 1982 to 2012. Sci Total Environ 2016;563–564:210–20.
[20] Beard JS. The vegetation survey of western Australia. Vegetatio 1975;30:179–87.
[21] Cawsey E, Austin M, Baker BL. Regional vegetation mapping in Australia: a case study in the practical use of statistical modelling. Biodivers Conserv 2002;11:2239–74.
[22] Naidoo R, Hill K. Emergence of indigenous vegetation classifications through integration of traditional ecological knowledge and remote sensing analyses. Environ Manage 2006;38:377–87.
[23] Molnár Z, Bartha S, Seregélyes T, et al. A grid-based, satellite-image supported, multi-attributed vegetation mapping method (méta). Folia Geobot 2007;42:225–47.
[24] Pedrotti F. Plant and vegetation mapping. Berlin: Springer Science & Business Media; 2012.
[25] Leprieur C, Kerr Y, Mastorchio S, et al. Monitoring vegetation cover across semi-arid regions: comparison of remote observations from various scales. Int J Remote Sens 2000;21:281–300.
[26] Xie Y, Sha Z, Yu M. Remote sensing imagery in vegetation mapping: a review. J Plant Ecol 2008;1:9–23.
[27] Roy PS, Behera MD, Murthy M, et al. New vegetation type map of India prepared using satellite remote sensing: comparison with global vegetation maps and utilities. Int J Appl Earth Obs 2015;39:142–59.
[28] Lawley V, Lewis M, Clarke K, et al. Site-based and remote sensing methods for monitoring indicators of vegetation condition: an Australian review. Ecol Indic 2016;60:1273–83.
[29] Xue J, Su B. Significant remote sensing vegetation indices: a review of developments and applications. J Sensors 2017:1353691.
[30] Hansen M, Coles J, Thomas KA, et al. USGS-NPS National Vegetation Mapping Program: Sunset Crater Volcano National Monument, Arizona, vegetation classification and distribution. US Geological Survey Final Project Report US Geological Survey (USGS): Reston, VA, 2004..
[31] Colgan M, Baldeck C, Féret JB, et al. Mapping savanna tree species at ecosystem scales using support vector machine classification and BRDF correction on airborne hyperspectral and LiDAR data. Remote Sens-Basel 2012;4:3462–80.
[32] Su Y, Guo Q, Fry DL, et al. A vegetation mapping strategy for conifer forests by combining airborne LiDAR data and aerial imagery. Can J Remote Sens 2016;42:1–15.
[33] Aschbacher J, Ofren R, Delsol JP, et al. An integrated comparative approach to mangrove vegetation mapping using advanced remote sensing and GIS technologies: preliminary results. Hydrobiologia 1995;295:285–94.
[34] Mitri GH, Gitas IZ. Mapping postfire vegetation recovery using EO-1 Hyperion imagery. IEEE Trans Geosci Remote Sensing 2009;48:1613–8.
[35] Mehner H, Cutler M, Fairbairn D, et al. Remote sensing of upland vegetation: the potential of high spatial resolution satellite sensors. Glob Ecol Biogeogr 2004;13:359–69.
[36] Mallinis G, Koutsias N, Tsakiri-Strati M, et al. Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. ISPRS-J Photogramm Remote Sens 2008;63:237–50.
[37] Bhaskaran S, Paramananda S, Ramnarayan M. Per-pixel and object-oriented classification methods for mapping urban features using IKONOS satellite data. Appl Geogr 2010;30:650–65.
[38] Jebur MN, Mohd Shafri HZ, Pradhan B, et al. Per-pixel and object-oriented classification methods for mapping urban land cover extraction using SPOT 5 imagery. Geocarto Int 2014;29:792–806.
[39] Trisurat Y, Eiumnoh A, Murai S, et al. Improvement of tropical vegetation mapping using a remote sensing technique: a case of Khao Yai National Park. Thailand. Int J Remote Sens 2000;21:2031–42.
[40] Cingolani AM, Renison D, Zak MR, et al. Mapping vegetation in a heterogeneous mountain rangeland using Landsat data: an alternative method to define and classify land-cover units. Remote Sens Environ 2004;92:84–97.
[41] Chen B, Xu B, Zhu Z, et al. 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. Sci Bull 2019;64:370–3.
[42] Gong P, Chen B, Li X, et al. Mapping essential urban land use categories in China (EULUC-China): preliminary results for 2018. Sci Bull 2020;65:182–7.
[43] Dymond CC, Mladenoff DJ, Radeloff VC. Phenological differences in tasseled cap indices improve deciduous forest classification. Remote Sens Environ 2002;80:460–72.
[44] Tigges J, Lakes T, Hostert P. Urban vegetation classification: benefits of multitemporal rapideye satellite data. Remote Sens Environ 2013;136:66–75.
[45] Huang B, Zhang H, Song H, et al. Unified fusion of remote-sensing imagery: generating simultaneously high-resolution synthetic spatial-temporal-spectral Earth observations. Remote Sens Lett 2013;4:561–9.
[46] Chen J, Chen J, Liao A, et al. Global land cover mapping at 30 m resolution: a POK-based operational approach. ISPRS-J Photogramm Remote Sens 2015;103:7–27.
[47] Chen B, Huang B, Xu B. Comparison of spatiotemporal fusion models: a review. Remote Sens-Basel 2015;7:1798–835.
[48] Liang L, Schwartz MD, Wang Z, et al. A cross comparison of spatiotemporally enhanced springtime phenological measurements from satellites and ground in a northern US mixed forest. IEEE Trans Geosci Remote Sensing 2014;52:7513–26.
[49] Zheng Y, Wu B, Zhang M, et al. Crop phenology detection using high spatio-temporal resolution data fused from SPOT5 and MODIS products. Sens-Basel 2016;16:2099.
[50] Didan K, Munoz BA, Solano R, et al. MODIS vegetation index user's guide (MOD13 series), version 3.0, June 2015 (Collection 6). 2015. Available online: https://vip.arizona.edu/documents/MODIS/MODIS_VI_UsersGuide_June_2015_C6.pdf (accessed on 27 March 2020)..
[51] Sulla-Menashe D, Gray JM, Abercrombie SP, et al. Hierarchical mapping of annual global land cover 2001 to present: the MODIS collection 6 land cover product. Remote Sens Environ 2019;222:183–94.
[52] Zhao X, Su Y, Hu T, et al. A global corrected SRTM DEM product for vegetated areas. Remote Sens Lett 2018;9:393–402.
[53] Farr TG, Kobrick M. Shuttle Radar Topography Mission produces a wealth of data. AGU Eos 2000;81:583–5.
[54] Rabus B, Eineder M, Roth A, et al. The Shuttle Radar Topography Mission—a new class of digital elevation models acquired by spaceborne radar. ISPRS-J Photogramm Remote Sens 2003;57:241–62.
[55] Su Y, Guo Q. A practical method for SRTM DEM correction over vegetated mountain areas. ISPRS-J Photogramm Remote Sens 2014;87:216–28.
[56] Su Y, Guo Q, Ma Q, et al. SRTM DEM correction in vegetated mountain areas through the integration of spaceborne LiDAR, airborne LiDAR, and optical imagery. Remote Sens-Basel 2015;7:11202–25.
[57] Fick SE, Hijmans RJ. Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol 2017;37:4302–15.
[58] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12:2825–30.
[59] Wang Z, Deng X, Song W, et al. What is the main cause of grassland degradation? A case study of grassland ecosystem service in the middle-south Inner Mongolia. Catena 2017;150:100–7.
[60] Chen C, Park T, Wang X, et al. China and India lead in greening of the world through land-use management. Nat Sustain 2019;2:122.

[61] Cui X, Graf HF. Recent land cover changes on the Tibetan Plateau: a review. Clim Change 2009;94:47–61.
[62] Du M, Kawashima S, Yonemura S, et al. Mutual influence between human activities and climate change in the Tibetan Plateau during recent years. Glob Planet Change 2004;41:241–9.
[63] Kussul N, Lavreniuk M, Skakun S, et al. Deep learning classification of land cover and crop types using remote sensing data. IEEE Geosci Remote Sens Lett 2017;14:778–82.
[64] Medina Machín A, Marcello J, Hernández-Cordero AI, et al. Vegetation species mapping in a coastal-dune ecosystem using high resolution satellite imagery. GISci Remote Sens 2019;56:210–32.
[65] Nogueira K, Dos Santos JA, Fornazari T, et al. Towards vegetation species discrimination by using data-driven descriptors. In: IAPR Workshop on Pattern Recognition in Remote Sensing. New York: IEEE, 2016.
[66] Guo K, Liu CC, Xie ZQ, et al. China vegetation classification: concept, approach and applications. Phytocoenologia 2018;48:113–20.

Yanjun Su is a professor at the Institute of Botany, Chinese Academy of Sciences. His research focuses on applying remote sensing to map vegetation structure and function parameters, and therefore understand the influence of anthropogenic activities and global climate change on terrestrial ecosystems.

Qinghua Guo is a professor at the Institute of Botany, Chinese Academy of Sciences. His research interests lie in the development of state-of-the-art geospatial techniques such as unmanned aerial vehicle, light detection and ranging, and big spatial data analytics to obtain high precision vegetation parameters for a better understanding of ecosystem patterns and functions from plot to regional and global scales.