An improved global land cover mapping in 2015 with 30 m resolution (GLC-2015) based on a multi-source product fusion approach

Bingjie Li [1], Xiaocong Xu [1], Xiaoping Liu [1][2], Qian Shi [1], Haoming Zhuang [1], Yaotong Cai [1] and Da He [1]

[1]School of Geography and Planning, Sun Yat-Sen University, Guangzhou, 510275, China

[2]Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, 519080, China

Correspondence to: Xiaoping Liu (liuxp3@mail.sysu.edu.cn)

Abstract. Global land cover (GLC) information with fine spatial resolution is a fundamental data input for studies on biogeochemical cycles of the Earth system and global climate change. Although there are several public GLC products with 30 m resolution, considerable inconsistencies were found among them especially in fragmented regions and transition zones, which brings great uncertainties to various application tasks. In this paper, we developed an improved global land cover map in 2015 with 30 m resolution (GLC-2015) by fusing multiple existing land cover products based on the Dempster-Shafer theory of evidence (DSET). Firstly, we used more than 160,000 global point-based samples to locally evaluated the reliability of the input GLC products for each LC class within each 4°×4° geographical grid for the establishment of the basic probability assignment (BPA) function. Then, the Dempster's rule of combination was used for each 30 m pixel to derive the combined probability mass of each possible land cover class from all the candidate maps. Finally, each pixel was determined with a land cover class based on a decision rule. Through this fusing process, each pixel is expected to be assigned with the land cover class that contributes to achieve a higher accuracy. We assessed our product separately with 34,711 global point-based samples and 201 global patch-based samples. Results show that, the GLC-2015 map achieved the highest mapping performance globally, continentally, and eco-regionally compared with the existing 30 m GLC maps, with an overall accuracy of 79.5% (83.6%) and a kappa coefficient of 0.757 (0.566) against the point-based (patch-based) validation samples. Additionally, we found that the GLC-2015 map showed substantial outperformance in the areas of inconsistency, with an accuracy

27 improvement of 19.3%-28.0% in areas of moderate inconsistency, and 27.5%-29.7% in areas of high

28 inconsistency. Hopefully, this improved GLC-2015 product can be applied to reduce uncertainties in the

29 research on global environmental changes, ecosystem service assessments, and hazard damage

30 evaluations, etc. The GLC-2015 map developed in this study is available at

31 https://doi.org/10.6084/m9.figshare.22358143.v2 (Li et al., 2022).

32 **1. Introduction**

33 Land cover (LC), influenced by both nature and human activities (Running, 2008; Gong et al., 2013;

34 Song et al., 2018; Liu et al., 2021a), is a significant component of the Earth system (Yang and Huang,

35 2021). Global land cover (GLC) products can serve as fundamental data for various studies, such as

36 climate and environmental changes (Bounoua et al., 2002; Foley et al., 2005; Grimm et al., 2008; Yang

37 et al., 2013; Schewe et al., 2019), food security (Verburg et al., 2013; Ban et al., 2015), carbon cycling

38 (Moody and Woodcock, 1994; Defries et al., 2002; Gómez et al., 2016), biodiversity conservation

39 (Chapin et al., 2000; Giri et al., 2005) and land management (Mayaux et al., 2004; Verburg et al., 2011).

40 Therefore, there is a pressing need for detailed, accurate, and high-quality GLC product to support global

41 change research and sustainable development.

42   In the preliminary stage, LC mapping mainly relied on visual interpretation, which is time-

43 consuming, labor-intensive and difficult to be applied at the global scale (Gong, 2012). In recent decades,

44 satellite remote sensing data, which can provide information of large area coverage and long-term

45 monitoring, has been adopted to generate GLC products. With coarse resolution satellite data such as

46 Advanced Very High Resolution Radiometer (AVHRR), Moderate Resolution Imaging

47 Spectroradiometer (MODIS), Medium Resolution Imaging Spectrometer (MERIS), and Global Land

48 Surface Satellite (GLASS), a variety of GLC products have been developed at 5 km to 300 m

49 resolution(Loveland et al., 2000; Hansen et al., 2000; Bartholomé and Belward, 2005; Friedl et al., 2010;

50 Defourny et al., 2018; Liu et al., 2020a). Although these GLC products have been widely applied to many

51 applications, it has been proved that the differences between sensors, classification systems, and

52 considerably low accuracies in areas prevent harmonization of these products (Herold et al., 2008;

53 Verburg et al., 2011; Grekousis et al., 2015). Also, these products are far from providing enough fine

54 spatial details of LC due to their relatively coarse spatial resolution, which does not meet the demand of

many studies (Giri et al., 2013; Yang et al., 2017). To allow researches which can capture most human activity, finer-resolution (e.g., 30 m) GLC products are demanded (Giri et al., 2013).

With the free accessibility of high-resolution satellite remote sensing data, GLC mapping at fine resolution has been successfully conducted. Using Landsat imagery, there has been a milestone achievement that the two GLC products are generated with fine resolution of 30 m, namely Finer Resolution Observation and Monitoring of Global Land Cover product (FROM_GLC)(Gong et al., 2013)and Globeland30 (Chen et al., 2015). After that, a 30 m-resolution GLC mapping in 2017 was achieved using the first all-season sample set (Li et al., 2017). More recently, Zhang et al. (2021) used both Landsat time series imagery and high-quality training data from the Global Spatial Temporal Spectra Library (GSPECLib) to produce a 30 m GLC map in 2015 (GLC_FCS30) with a two-level classification scheme. Several attempts have been made to improve accuracy of 30 m GLC products which are prevail in the generation of GLC mapping task over the last few years. FROM_GLC was created by employing four classification algorithms to classify the Landsat images and choosing time series of MODIS EVI data for training and test. Globeland30 was created by proposing a pixel-object-knowledge-based (POK) method to assure consistency and accuracy. GLC_FCS30 was generated by adopting local adaptive random forest models with high-quality training samples derived from GSPECLib. The Globeland30, FROM_GLC, and GLC_FCS30 are excellent and indispensable GLC products which have contributed much to various researches, such as biodiversity conservation (Wu et al., 2020; Meng et al., 2023), climate change (Kim et al., 2016; Xue et al., 2021; Zheng et al., 2022), and land management (Shafizadeh-Moghadam et al., 2019).

Despite the great efforts in producing more accurate products, the existing 30 m GLC products still show unstable performance in certain LC classes and some specific areas (Sun et al., 2016; Kang et al., 2020). Furthermore, the existing 30 m products showed great agreement in overall spatial distribution patterns but significant spatial inconsistency in some specific areas (heterogeneous areas and transition zones) and spectrally similar classes (forest and shrubland, cropland and grassland) (Gao et al., 2020; Liu et al., 2021b). The spatial inconsistency between the existing 30m GLC products are resulted from differences in their classification systems, classification techniques employed, source data, and spatial distribution and size of training samples (Yang et al., 2017; Gao et al., 2020). Due to the aforesaid limitations, users of GLC products still have difficulties in an appropriate selection of data for their

84   specific application. Ultimately, this situation leads to uncertainties in outcomes of related researches

85   when different 30 m GLC products are used. For GLC mapping with fine spatial resolution, more efforts

86   should be focused on improving the mapping in heterogenous and fragmented landscape (Herold et al.,

87   2008; Liu et al., 2021b). Therefore, it is pressing to generate a more accurate and reliable GLC product

88   with high classification accuracy, especially for spatially inconsistent regions and low-accuracy LC

89   classes.

90   According to Gong et al. (2016), inconsistencies between LC products indicate available

91   complementary information and more robust and reliable data can be generated by integrating the input

92   maps with the data fusion method. Given that different maps have disagreement and provide accurate

93   information in different locations, we can make a best choice for the class label assigned to each pixel

94   by weighting the credibility of all the available information and combining them through a decision rule

95   (Clinton et al., 2015). In this way, the output map of integration on input maps can reduce the overall risk

96   of assigning a wrong class label to a pixel and at least achieve the average performance of input maps.

97   Several attempts have been made to produce an accurate and consistent LC map using various methods,

98   such as majority voting (MV), fuzzy agreement and Bayesian theory. Iwao et al. (2011) created a GLC

99   map based on a simple majority voting method. Jung et al. (2006) generated a 1km GLC map by

100   combination of MODIS, GLC2000 and GLCC data based on fuzzy agreement scoring. Subsequently,

101   Fritz et al. (2011) extended the synergy method of Jung et al. (2006) by ranking LC maps and mapped

102   the cropland extent in Sub-Saharan Africa. See et al. (2015) generated two GLC products by integrating

103   medium resolution LC products with geographically weighted regression (GWR). Gengler and Bogaert

104   (2018) proposed a Bayesian data fusion method and applied it to the LC mapping for a specific region in

105   Belgium. All these researches have demonstrated that fusion method can create an integrated LC product

106   where the mapping accuracy is greatly improved by combing the best of candidate maps. However, the

107   MV method is sensitive to the quality of the candidate maps and has significant uncertainties when the

108   input products exhibit great disagreement(Chen and Venkataramanan, 2005). The fuzzy agreement is

109   highly subjective since it depends on expert assessment, while the Bayesian theory requires a prior

110   knowledge or conditional probabilities and fails to handle the states of ignorance(Liu and Xu, 2021).

111   The Dempster-Shafer theory of evidence (DSET) is an evidence-based approach to reason with

112   uncertainties. Unlike the majority voting, the DSET method can discount evidence from inaccurate

information with a probability mass that reflects the degree of belief rather than a binary decision (Razi et al., 2019). In contrast to the Bayesian theory, the DSET can integrate evidence from a variety of sources without the requirement of prior knowledge (Chen and Venkataramanan, 2005). Moreover, the reliability of the final fused results is measured the DSET method with a total degree of belief. Although previous literature focused on the application of the DSET method in multisource data aggregation, very little research has been conducted at a global scale due to the lack of accurate and sufficient samples and the demand for adequate computing resources.

In this research, we propose a multi-source product fusion approach on the Google Earth Engine (GEE) platform to produce an improved GLC product in 2015 (GLC-2015) with 30 m resolution. The fusion approach we proposed aims to deal with the inconsistency between previous 30 m GLC products and generate a map which has better mapping performance than any of the candidate maps by evaluating the mapping accuracy of these existing products at the local scale and choosing the most credible LC class. To fulfill the purpose, we first performed reliability evaluation, where the accuracy of each GLC product for each LC class in each $4° \times 4°$ geographical grid is regarded as the evidential probability to create the basic probability assignment (BPA) function. Then, the BPA values of all the LC classes from different GLC products are fused according to the Dempster's rule of combination. Finally, the GLC-2015 map was integrated after a final accepted LC class with the maximum combined probability mass was assigned to each 30 m pixel. Our GLC-2015 map was separately validated with two different validation sets, namely global point-based samples and global patch-based samples, and compared with three existing multiple-class GLC products. Moreover, we provided an analysis for mapping improvement of the GLC-2015 compared to other products in areas of high mapping inconsistency. The GLC-2015 map is proved to be accurate and credible and can significantly improve the mapping accuracy in areas of high inconsistency between previous products.

2. Datasets

2.1 Multiple-class GLC products

Three existing 30m GLC products with multiple classes, including GlobeLand30, FROM_GLC and GLC_FCS30, were employed as input maps in the fusion based on DSET. A summary of their detailed information is shown in Table 1.

GlobeLand30, a widely-used global geo-information product, was produced by the POK-based method using Landsat and HJ-1 satellite images. Globeland30 products are freely accessible online at the website (http://www.globalland30.org) for 2000 and 2010. From the accuracy assessment, the Globeland30 for the year 2010 had an overall accuracy excessed 80% using large samples (Chen et al., 2015). Although the data time of GlobeLand30 is 2010, which has a five-year gap with other products, it was used in our project for the following reasons: (1) The changed areas of LC caused by the time interval are tiny compared to the global land area. In addition, there is relatively less uncertainty due to LC changes than due to inaccurate classification (Xu et al., 2014). Most spatial disagreements between the existing maps are about classification errors rather than LC changes over the time interval (Mccallum et al., 2006; See et al., 2015); (2) We used a global point-based sample set for the year 2015 to evaluate the reliability of the input products in all 4° × 4° grids. At locations where land cover changed between 2010 and 2015, the Globeland30 was more likely to have low accuracy based on the validation and less likely to contribute to the fusion using the DSET approach. In this way, the errors due to land cover changes can be largely avoided.

FROM_GLC was first generated using numerous Landsat images, which has a fine classification system with a two-level structure. It achieved an OA of 64.5% through validation with the complete test samples and 71.5% with a subset of test samples in homogeneous areas (Gong et al., 2013). We used the version of 2015 for the fusion.

GLC_FCS30 was developed using Landsat time series data and large training samples from the GSPECLib. It has a two-level classification scheme that contains 16 global LCCS LC classes and 14 detailed regional LC classes. The overall accuracy of the GLC_FCS30 according to LCCS level-1 validation scheme reached 71.4% (Zhang et al., 2021).

Land cover products which focus on a national scale are more likely to possess higher accuracy because they were produced by experts who have good knowledge of land cover classes nationally. Thus, the National Land Cover Database 2016 (NLCD 2016) for the year 2016 (Yang et al., 2018), China's land-use/cover datasets (CLUDs) (Liu et al., 2014) for 2015, and the annual China land cover dataset (CLCD) (Yang and Huang, 2021) for 2015 were also included in the fusion. NLCD 2016 database, which provides continuous and accurate information about land cover and change from 2001 to 2016 at an interval of 2 or 3 years, was produced based on a pixel- and object-based approach and an effective post-

170 classification process (Yang et al., 2018). The level-1 and level-2 overall accuracy of NLCD 2016

171 database for 2016 was 90.6% and 86.4%, respectively (Wickham et al., 2021). CLUDs, developed by the

172 digital interpretation method using Landsat images, provide land cover information over China from

173 1980s to 2015. The overall accuracy of CLUDs reached 94.3% and 91.2% for level-1 and level-2 land

174 cover classes, respectively (Liu et al., 2014). CLCD was generated with stable training samples derived

175 from CLUDs and Landsat time series. Assessed with 5463 validation samples, CLCD obtained an overall

176 accuracy of 79.31% (Yang and Huang, 2021).

177 **Table 1. Detailed information of GLC products and national-scale LC products used in this paper.**

Product name	Satellite sensors	Year of reference	Access	Literature
Globeland30	Landsat TM/ETM+ HJ-1 A/B	2010	http://www.globallandcover.com/	(Chen et al., 2015)
FROM_GLC	Landsat TM/ETM+/OLI	2015	http://data.ess.tsinghua.edu.cn/	(Gong et al., 2013)
GLC_FCS30	Landsat OLI	2015	https://doi.org/10.5281/zenodo.3986872	(Zhang et al., 2021)
GAUD	Landsat TM/ETM+/OLI	2015	https://doi.org/10.6084/m9.figshare.115131 78.v1	(Liu et al., 2020b)
GFC	Landsat TM/ETM+	2015	http://earthenginepartners.appspot.com/scie nce-2013-global-forest	(Hansen et al., 2013)
JRC GSW	Landsat TM/ETM+/OLI	2015	http://global-surface-water.appspot.com/	(Pekel et al., 2016)
GMW	ALOS PALSAR Landsat TM/ETM+	2015	https://data.unep-wcmc.org/datasets/45	(Bunting et al., 2018)
NLCD 2016	Landsat TM /OLI	2016	https://www.mrlc.gov/data/nlcd-2016-land- cover-conus	(Yang et al., 2018)
CLUDs	Landsat TM HJ-1 CBERS-1	2015	/	(Liu et al., 2014)
CLCD	Landsat TM/ETM+/OLI	2015	https://doi.org/10.5281/zenodo.4417810	(Yang and Huang, 2021)

178 **2.2 Single-class GLC products**

179 To improve the quality of the fusing result, a set of highly qualified GLC products with single class at 30

180   m fine resolution were also used. Compared to the multiple-class GLC products, these single-class GLC

181   products are more likely to provide accurate information since they usually focus on promoting mapping

182   performance of specific LC class. These products include Global Forest Change (GFC) (Hansen et al.,

183   2013), Global Annual Urban Dynamics (GAUD) (Liu et al., 2020b), Joint Research Centre's Global

184   Surface Water (JRC GSW) (Pekel et al., 2016), and Global Mangrove Watch (GMW) (Bunting et al.,

185   2018). While these single-class products are either annual or multi-epoch, we only selected these

186   products in the target year of 2015. The background information of these single-class products was

187   considered as another land cover class (e.g., non-water) participating in the fusion. The accuracy of the

188   background information was defaulted to 0 since it did not provide information about any of the other

189   nine categories in our classification system. Table 1 also describes the information of these selected

190   single-class GLC products.

191   GFC was resulted from a time-series analysis of growing season Landsat scenes, aiming to provide

192   information about global tree cover extent, gain, and loss at a 30m spatial resolution. The accuracy

193   assessment was performed at global and climate domain scales and the forest gain reached an overall

194   accuracy of 99.6% and forest loss reached 99.7% across the globe (Hansen et al., 2013). Up to now, it

195   has a temporary coverage from 2000 to 2020.

196   GAUD, which provides 30m annual urban extent for the time period of 1985 to 2015, was generated

197   using numerous Landsat images with both data fusion approach and temporal segmentation approach on

198   the GEE platform. Validation was conducted across different urban ecoregions and the globe by the

199   product developer. The accuracy of mapping urbanized year was 76% for the period of 1985 to 2000 and

200   82% for the period of 2000 to 2015 at humid regions worldwide (Liu et al., 2020b).

201   JRC GSW dataset provides a monthly presentation of global surface water changes from 1984 to

202   2015 at a fine 30 m resolution. Expert systems, visual analytics and evidential reasoning were exploited

203   to detect water extent and changes. Based on 40,124 validation points over the globe and across the 32

204   years, commission accuracies were determined with overall accuracies of 99.45% (TM), 99.35% (ETM+)

205   and 99.54% (OLI) and omission accuracies were reflected in overall accuracies of 97.01% (TM), 95.79%

206   (ETM+) and 96.25%(OLI) (Pekel et al., 2016). We used the GSW Yearly Water Classification History

207   v1.1 in the GEE catalog. A single 'waterClass' band is present in each image that provides the water's

208   seasonality throughout the year with four types: no data, no water, seasonal water, and permanent water.

209 Since the seasonal water in GSW data is not as reliable as the permanent water (Meyer et al., 2020), we

210 selected permanent water bodies and excluded seasonal water bodies.

211 GMW dataset was produced as a resulted of the GMW initiative, which aims to provide consistent

212 information of mangrove extent. The global mangrove map in 2010 was generated as a baseline map

213 employing the Extremely Randomized Trees classifier to classify ALOS PALSAR and Landsat imagery.

214 Assessed by a total of 53,878 sample points globally, the overall accuracy of the baseline map reached

215 95.3% and the producer's accuracy achieved 94.0% (Bunting et al., 2018). Based on the baseline in 2010,

216 mangrove extent maps for six epochs between 1996 and 2016 have been established and annual change

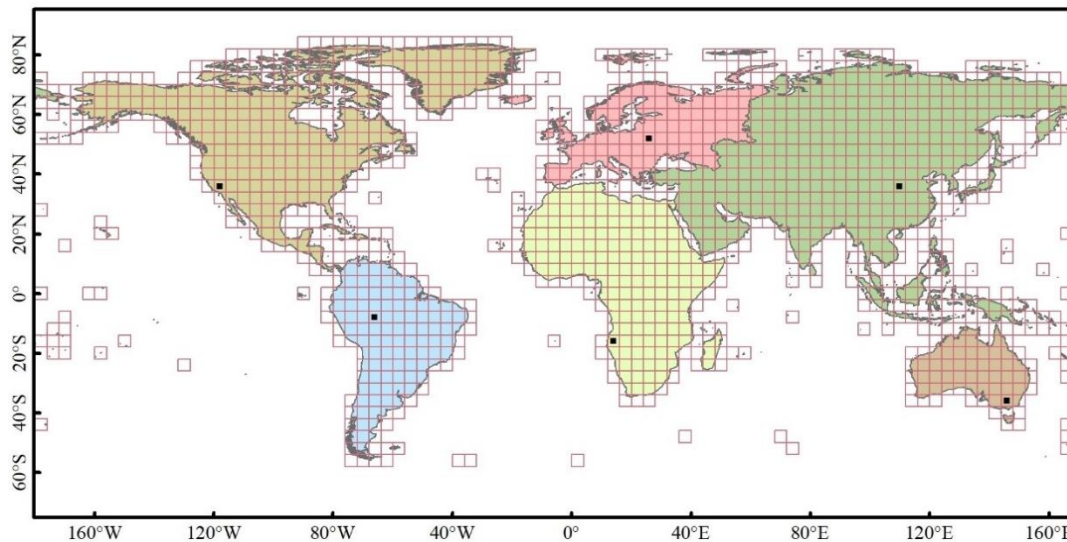
217 monitoring from 2018 and onwards are undertaken.

218 **2.3 Global point-based and patch-based samples**

219 In this study, we collected two sets of global samples, namely the global point-based samples and the

220 global patch-based samples. To collect representative and sufficient samples efficiently, we divided the

221 world's terrestrial area into 4° × 4° geographical grids. A total of 1,507 grids are distributed evenly across

222 the globe, shown as Figure 1.

224 **Figure 1. Spatial distribution of the 4° × 4° geographical grids over the world. Six black rectangle tiles with**
225 **size of 0.25° were used for visual comparation between our product and other three products.**

226 To derive the global point-based samples, we adopted stratified random sampling in each grid. The

227 stratified random sampling depends on area ratio of LC classes from a LC product. We used the

228 FROM_GLC as prior knowledge rather than the Globeland30 and GLC_FCS30 with two considerations:

229   (1) the FROM_GLC has the same data time as our target map (GLC-2015) while the Globeland30 has a

230   5-year interval from our samples, which affects the size of samples for each LC class; (2)the 10 level-1

231   land cover classes of the FROM_GLC is similar to that in the classification system of the GLC-2015,

232   while the GLC_FCS30 has differences with the GLC-2015 in the classification scheme and definition of

233   land cover classes. First, the FROM_GLC product was used to calculate the area ratio of each LC class.

234   Then, points were randomly extracted from the FROM_GLC according to the area ratio and spatial

235   location of each class. Finally, more than 200,000 global samples were collected. Through the sampling

236   method mentioned above, the global point-based samples were even across the globe and sufficient for

237   each LC class in each grid. Therefore, more than 50 points could be easily derived for LC classes with a

238   small area ratio in the $4° × 4°$ grid. The FROM_GLC shows low accuracy for some LC classes, especially

239   for cropland and forest (Gao et al., 2020; Liu et al., 2021b; Zhang et al., 2021; Zhang et al., 2022). If the

240   global samples were extracted with LC class label from the FROM_GLC, there would be inevitable

241   errors. Therefore, the FROM_GLC was only used to determine the size and location of samples for each

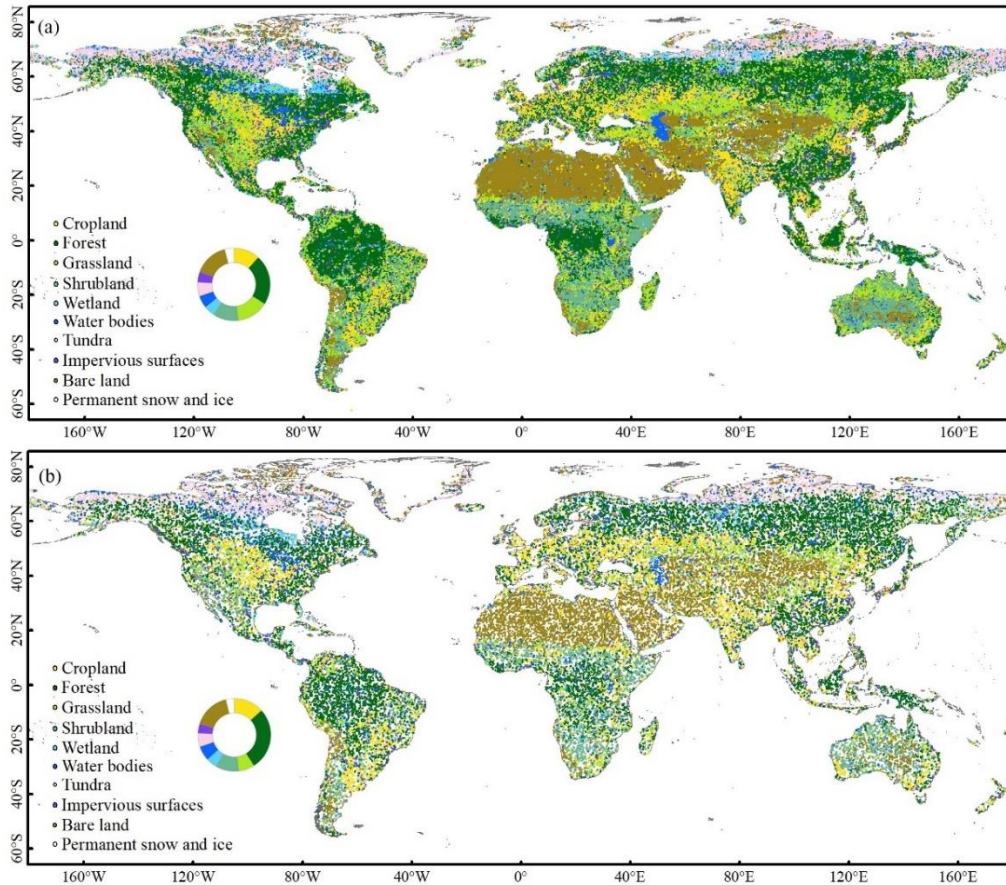
242   LC class. Instead, all the points were manually labeled according to Google Earth high-resolution images.

243   The whole sample set was randomly split into two subsets: 80% of the global samples were used to assess

244   the accuracy of each GLC product for various LC classes at the global scale and in each grid. The

245   remaining 20% were used for the validation of the GLC-2015 map and data inter-comparison between

246   different GLC products. Figure 2 presents the distribution of the whole global point-based samples and

247   the subset for accuracy assessment and data inter-comparison.

248
249 **Figure 2. Spatial distribution of (a) the global point-based samples, (b) the subset of the global point-based**
250 **samples for accuracy assessment and data inter-comparison, the proportions of each LC class are shown in**
251 **the pie chart.**

252   To verify the consistency between the GLC-2015 and the actual pattern of the landscape at the local

253 scale, we also established the global patch-based samples. Simple random sampling was used to derive

254 5 km × 5 km blocks over the world's terrestrial area and across different ecoregions because it is easy to

255 perform and capable to augment the sample size from target areas (Pengra et al., 2020). Since

256 inconsistency between current GLC maps tends to appear in those heterogeneous areas, such as

257 fragmented regions and transition zones, we slightly increased the sample size for areas with the

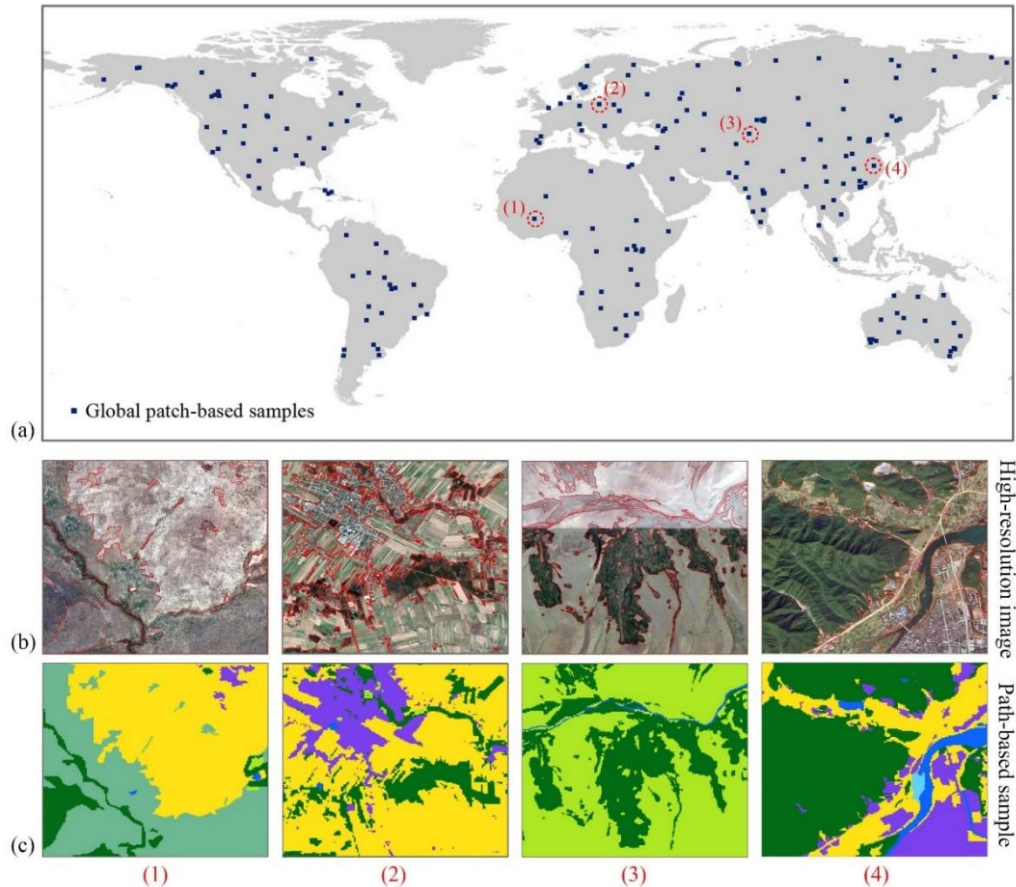
258 heterogeneous landscape to better evaluate our mapping results. In total, there were 201 blocks selected

259 as the global patch-based samples, as displayed in Figure 3a. Then, for each block in the patch-based

260 samples, we used ArcGIS 10.5 software to derive polygons (patches) of various sizes which captured the

261 real landscape on the high-resolution images. Meanwhile, each polygon was manually labeled with a LC

262 class. Four examples of producing patch-based samples are shown in Figure 3b, c.

263

264 **Figure 3. Spatial distribution and selected examples of the global patch-based samples. The location of 5 km**

265 **× 5 km patch-based samples are shown as panel (a), the locations of four selected samples are remarked by**

266 **red dash circles. Panels (b) and (c) illustrate the production of global patch-based samples on manual**

267 **interpretation. The red lines in high-resolution images circa 2015 are results after vectorization using ArcGIS**

268 **10.5 software. Four corresponding patch-based samples are shown as (c).**

269 **3.   Methods**

270   In this study, we proposed a multi-source product fusion method to produce the GLC-2015 map. The

271   procedure mainly comprised the fusion based on the Dempster-Shafer theory of evidence (DSET),

272   accuracy assessment and data inter-comparison (Figure 4). The basic of this study is the fusion of multi-

273   source GLC products based on DSET. The fusion method was performed at the pixel level and it involves

274   the following three main steps: (1) Construct the basic probability assignment (BPA) function of each

275   pixel that belongs to each LC class considering the accuracy assessment of different GLC products; (2)

276   calculate the combined probability mass for each class per pixel using the Dempster's rule of combination;

277   and (3) determine the finally accepted LC class per pixel by a decision rule. Afterwards, pixels with a

278   determined LC class were integrated to generate a new map. For large-scale or global land cover mapping,

279   previous researchers divided the study area into a lot of sub-regions and conducted classification in each

280   sub-region on GEE (Gong et al., 2020; Liu et al., 2020b; Huang et al., 2021; Jin et al., 2022; Zhang et

281   al., 2021; Zhao et al., 2021). The shape and size of sub-region vary in previous work, such as hexagons

282   with a side length of 2°, geographical grids with a size of 1°×1°, 3.5°×3.5°, 5°×5°, or 10°×10°. When

283   deciding on the size of sub-regions, two important factors should be considered. The size of samples in

284   each sub-region should be sufficient so that the rare land cover classes will not be missed. On the other

285   hand, it is impossible to implement mapping work at a sub-region as larger as we want due to memory

286   constraints. To determine the appropriate size, we tested different sizes of the sub-region (see Table S1).

287   Result shows that dividing the study area into 4°×4° grids performed best. Therefore, we split the world's

288   terrestrial area into 1507 4°×4° geographical grids. The entire framework was implemented in all 4° × 4°

289   geographical grids on the GEE platform.

290

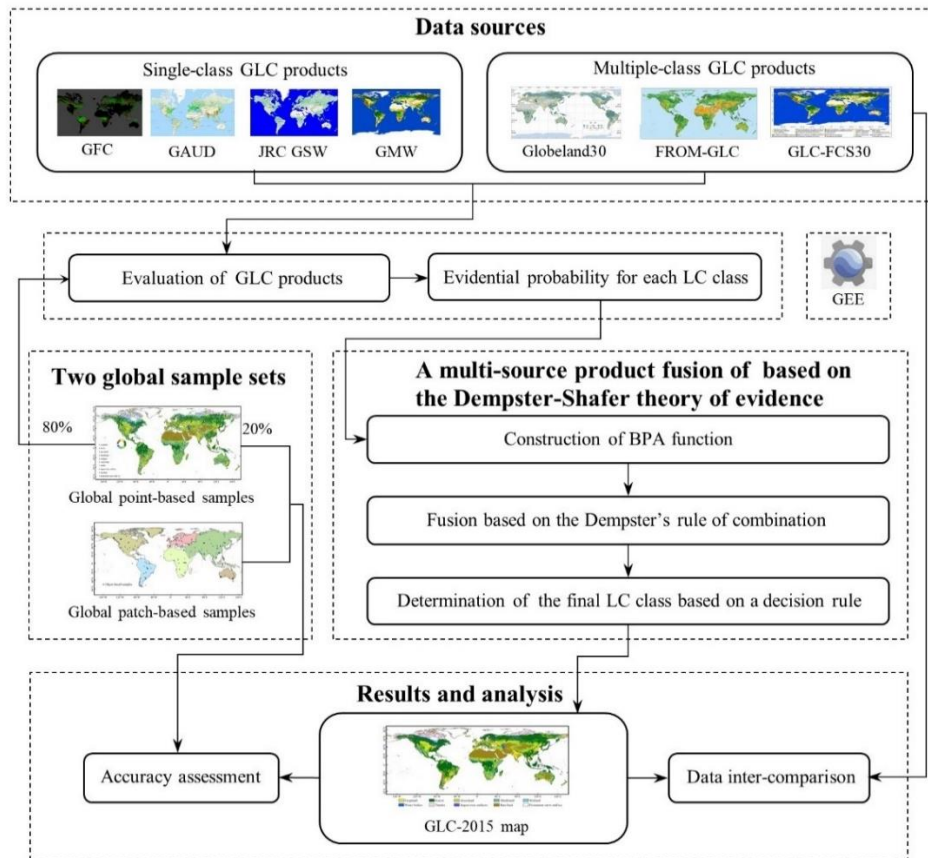



Figure 4. The framework for generating the GLC-2015 map using a multi-source product fusion approach
based on DEST.

3.1 Definition of the classification system

294   In this study, we adopted the classification system with 10 LC classes, including cropland, forest,

295 grassland, shrubland, wetland, water bodies, tundra, impervious surfaces, bare land, and permanent snow

296 and ice (Chen et al., 2015), as listed in Table 2. Due to the applications for different social needs, the

297 existing GLC products and national-scale LC products were produced with different classification

298 systems (Tables S2-S3). The GlobeLand30 used a simple classification system that only contained 10

299 first-level classes. Unlike the GlobeLand30, the FROM_GLC and GLC_FCS30 were classified with a

300 two-level classification scheme. Through analysis of these systems, we found that the classification

301 systems are not the same, but they have some agreements. There are both 10 major classes in the

302 GlobeLand30 and FROM_GLC despite that the definition of some classes differs. Additionally, in

303 contrast to the GlobeLand30 and FROM_GLC, the level-0 classification system of the GLC_FCS30

304 lacks tundra. However, in the level-2 detailed LC classes of the GLC_FCS30, lichens and mosses has

305 little distinction with tundra.

306 **Table 2. Classification system adopted in this paper.**

Id	LC class	Definition
10	Cropland	Land areas used for food production and animal feed.
20	Forest	Land areas dominated by trees with tree canopy cover over 30%, and sparse trees with tree canopy cover between 10%-30%.
30	Grassland	Land areas dominated by natural grass with a cover over 10%.
40	Shrubland	Land areas dominated by shrubs with a cover over 30%, including mountain shrubs, deciduous shrubs, evergreen shrubs and desert shrubs with a cover over 10%.
50	Wetland	Land areas dominated by wetland plants and water bodies.
60	Water bodies	Land areas covered with accumulated liquid water.
70	Tundra	Land areas dominated by lichen, moss, hardly perennial herb and shrubs in the polar regions.
80	Impervious surfaces	Land areas covered with artificial structures.
90	Bare land	Land areas with scarce vegetation with a cover lower than 10%.
100	Permanent snow and ice	Land areas dominated by permanent snow, glacier and icecap.

307 According to the classification system adopted in the study, the original LC classes of FROM_GLC

308 and GLC_FCS30, CLUD for 2015, and NLCD 2016 for 2016 were converted into the 10 target land

309 cover classes based on the similarity of LC definition. Note that cropland in our classification system

310 was defined as land areas for food production and animal feed. Therefore, pasture in level-2 classes of

313 **3.2 A multi-source product fusion for the GLC-2015 mapping**

314 The DSET is an effective method widely applied for the fusion of multi-source data. To generate a new

315 high-quality GLC map, a multi-source product fusion method using DSET was proposed. In the

316 remainder of the section 3.2, We introduced the overview on the theory and presented the application of

317 DSET in our mapping process.

318 **3.2.1 Dempster-Shafer theory of evidence**

319 The DSET is developed by Dempster and Shafer, which is an extension of Bayesian probability theory.

320 This theory treats information from different data sources as independent evidence and integrated these

321 evidences with no requirements regarding the prior knowledge. In the fusion, we assume a classification

322 process in which all the input data are to be classified into mutually exclusive classes. Let the set Ω of

323 these classes be a frame of discrimination. 2^{Ω} is the power set of Ω that includes all the classes and

324 their possible unions. We defined the function $m: 2^{\Omega} \to [0,1]$ as the basic probability assignment (BPA)

325 function if and only if it satisfies $m(\Phi) = 0$ and $\sum_{A \subseteq 2^{\Omega}} m(A) = 1$ with Ø denotes an empty set. For

326 each class $A \subseteq 2^{\Omega}$, m(A) is called the basic probability mass which can be computed from the BPA

327 function and represents the degree of support for class A or confidence in class A.

328   The purpose of fusion is to evaluate and integrate information from multiple sources. In the DSET,

329 these multi-source data are regarded as different evidence and provide different assessments. To generate

330 all the evidences, Dempster-Shafer theory of evidence offers a rule. Suppose $m_i(B_j)$ is the basic

331 probability mass computed from the BPA function for each input data i with $1 \le i \le n$ for all classes

332 $B_j \in 2^{\Omega}$. Dempster's rule of combination is provided to calculate a combined probability mass from

333 different evidences. The fusion rules are given in equation (1) and (2).

$$m(C) = \frac{\sum_{B_1 \cap B_2 \ldots \cap B_n = C} \prod_{1 \le i \le n} m_i(B_j)}{1 - k} \tag{1}$$

334

335

$$k = \sum_{B_1 \cap B_2 \ldots \cap B_n = \emptyset} \prod_{1 \le i \le n} m_i(B_j) \tag{2}$$

336

337 Where k represents the basic probability mass associated with conflicts among the sources of evidence.

338   C is the intersection of all classes B_j and carries the joint information from all the input data. After the

339   combination, we took a decision rule to decide the class we finally accept. There are several ways to

340   decide the final class by simply choosing the class with the maximum belief, plausibility, support, or

341   commonality.

342   **3.2.2 Mapping based on DSET**

343   Here, we presented our implementation for the GLC-2015 mapping in the framework of DSET. All the

344   multiple-class and single-class GLC products described in Sect. 2 were selected as input maps to be

345   combined. In the integration of multi-source GLC products, since all the LC classes in our classification

346   system are known, the frame of discrimination was defined to be our classification system:

347    $$\Omega = \begin{Bmatrix} \text{cropland, forest, grassland, shrubland, wetland, water bodies,} \\ \text{tundra, impervious surfaces, bare land, permanent snow and ice} \end{Bmatrix} \tag{3}$$

348   The definition of BPA function is the critical point in applying DSET (Rottensteiner et al., 2005).

349   In the fusion, we wanted to achieve a per-pixel classification into one of ten LC classes: cropland, forest,

350   grassland, shrubland, wetland, water bodies, tundra, impervious surfaces, bare land, and permanent snow

351   and ice. For each single-class or multiple-class GLC product, the accuracy for each LC class was

352   calculated and used as evidential probability to construct the BPA. Given that the local accuracy for a

353   4°×4° grid was not able to adequately reflect the actual land cover landscape, especially for the rare LC

354   classes, the global accuracy was incorporated into the construction of the BPA to avoid uncertainties from

355   a local point of view. Since the assessment based on local samples plays a more critical role in BPA

356   construction for a local grid, a higher weight should be assigned to the local accuracy. To identify the

357   best weight, we tested different weights of the local accuracy (see Figure S1). The esult shows that using

358   75% performed robustly and obtained relatively higher overall accuracy. Therefore, we chose 75% as

359   the weight for local accuracy and 25% for global accuracy. Here, we defined the BPA function as follow:

360    $$m_i(T_j) = \frac{PA_{local_{(ij)}} + UA_{local_{(ij)}}}{2} \times 75\% + \frac{PA_{global_{(ij)}} + UA_{global_{(ij)}}}{2} \times 25\% \tag{4}$$

361   Where $m_i(T_j)$ represents the BPA function of evidence source i for the LC class T_j ; $PA_{local_{(ij)}}$,

362    $UA_{local_{(ij)}}$ denote producer's accuracy and user's accuracy of evidence source i for the LC class T_j for

363   each $4° \times 4°$ geographical grid, respectively; $PA_{global_{(ij)}}$, $UA_{global_{(ij)}}$ denote producer's accuracy and

364   user's accuracy of evidence source i for LC class T_j at the global scale.

365   To estimate the exact values of $PA_{local_{(ij)}}$, $UA_{local_{(ij)}}$, $PA_{global_{(ij)}}$ and $UA_{global_{(ij)}}$, we used 80%

366   of the global point-based samples more than 160,000 points derived in Sect 2.3. As soon as we obtained

367   the measurements of $m_i(T_j)$, the combined probability masses $m(T_j)$ were evaluated based on

368   Dempster's rule of combination for each pixel classified as the LC class T_j by fusing BPA values of all

369   the evidence sources:

$$m(T_j) = \frac{1}{1-k} \sum_{T_{1j} \cap T_{2j} \dots \cap T_{nj} = T_j} m_i(T_j) \tag{5}$$

370

$$k = \sum_{T_{1j} \cap T_{2j} \dots \cap T_{nj} = \emptyset} m_i(T_j) \tag{6}$$

371

372   Where k represents the basic probability mass associated with conflict; $m_i(T_j)$ represents the basic

373   probability mass of a certain pixel belonging to the LC class T_j from different GLC products.

374   Additionally, a belief measure (Bel) was given to measure the degree of credibility that a pixel

375   labeled as the finally accepted LC class when combining all the available evidences. The belief measure

376   was determined by

$$Bel(T_j) = \sum_{T_{ij} \subseteq T_j} m_i(T_j) \tag{7}$$

377

378   To determine the finally accepted LC class per pixel, we took the rule of maximum combined

379   probability mass as our decision rule and the LC class with the maximum combined probability mass is

380   assigned to the 30 m pixel. Pixels labeled with the LC class were integrated to generate the GLC-2015

381   product.

382   **3.3 Accuracy assessment**

383   To assess the accuracy of the GLC-2015 map, we utilized two validation methods: validation with the

384   global point-based samples and the global patch-based samples. Since the global point-based sample set

385   is distributed evenly across the world and its sample size for each LC class is relatively sufficient and

386   balanced, even for the rare classes, it can provide a representative and credible basis for estimation of the

387   GLC-2015 map globally. Furthermore, we used the global patch-based samples to conduct accuracy

388   assessment from the local landscape scale. Although the global patch-based sample set provide an

389   inadequate sample size for rare LC classes, it can take advantage of the spatial context information and

390   efficiently reflect the actual pattern of the landscape.

391   The confusion matrix was produced to evaluate and analyze the GLC-2015 mapping result. The

392   error matrix is composed of entry A_{ij}, which represents the number of samples with reference LC class

393   j being classified as LC class i.The overall accuracy (OA), kappa coefficient, producer's accuracy (PA),

394   and user's accuracy (UA) were generated from confusion matrix to describe the quality of the GLC-2015

395   map. They are defined as follows:

$$OA = \frac{\sum_i A_{ii}}{\sum_i \sum_j A_{ij}} \tag{8}$$

$$P_o = OA \tag{9}$$

$$P_e = \sum_k \frac{\sum_i A_{ik}}{\sum_i \sum_j A_{ij}} \times \frac{\sum_j A_{kj}}{\sum_i \sum_j A_{ij}} \tag{10}$$

$$kappa = \frac{P_o - P_e}{1 - P_e} \tag{11}$$

$$PA^i = \frac{A_{ii}}{\sum_k A_{ki}} \tag{12}$$

$$UA^i = \frac{A_{ii}}{\sum_k A_{ik}} \tag{13}$$

402   Where UA^i and PA^i represent UA and PA of the LC i, respectively; P_o is the agreement between the

403   reference and the classified data; P_e is the hypothetical probability of chance agreement.

404   **3.4 Data inter-comparison**

405   To better reflect the quality of the GLC-2015 map, we intercompared the GLC-2015 map with LC

406   products at multiple scales. In the accuracy assessment of different products, two global validation sets

407   described earlier were employed.

408   To figure out whether the GLC-2015 map promotes accuracy in the areas with high classification

409   difficulty and how much the improvement is compared to the other products, we conducted the spatial

410   consistency analysis between the GlobeLand30, FROM_GLC, and GLC_FCS30 and compared the

411   mapping performance of the GLC-2015 with others in the areas of low inconsistency, moderate

412   inconsistency, and high inconsistency. To visually present the spatial consistency between three existing

413   GLC maps, we employed the spatial superposition method to obtain the spatial correspondence pixel-

414   by-pixel between different maps. Based on the times of all the GLC products agreed for the same LC

415   class, the degree of consistency for a pixel was identified as three levels with the agreement value equal

416   to 3, 2, or 1. The areas of low inconsistency were regarded as pixels that classified as the same LC class

417   in all three GLC maps (labeled as 3). The moderate inconsistency areas were regarded as pixels that were

418   consistent in only two GLC maps (labeled as 2). The high inconsistency areas were regarded as pixels

419   that were totally inconsistent in these three GLC maps (labeled as 1). For a visual comparison, all these

420 GLC maps were aggregated to 0.05°, in which the LC class with the largest proportion determined the

421 class in each 0.05° grid.

3.5 Assessment on mapping performance of DSET and other methods

423 In addition to inter-comparison between the GLC-2015 map and three existing GLC products, we

424 compared the DSET method with two existing commonly used fusion methods, including the majority

425 voting (MV) and spatial correspondence (SC) based on two global validation sets including 20% of the

426 global point-based samples and the whole global patch-based samples. MV is a fusion approach that

427 combines input maps and adopts the LC class favored by the majority of the candidate maps. In the MV

428 method, we compared the GlobeLand30, FROM_GLC, and GLC_FCS30 at each pixel and chose the

429 class that two or three LC products agreed for. For pixels where three LC products were different, the

430 LC class of the product with the highest accuracy was adopted. SC method produces an integrated land

431 cover map by selecting the LC class of the input map that has the highest spatial correspondence with

432 the reference data. In this study, 80% of the global point-based samples were used as the reference data

433 to obtain the SC map of each global LC product. If the class of a product agreed with that of the point-

434 based sample, a value equal to 1 was assigned to that sample. On the contrary, a value equal to 0 was

435 assigned to the sample if the class of the product differed from that of the sample. In each 4° × 4° grid,

436 we used the Kriging method to obtain spatial correspondence maps which have the correspondence value

437 ranging from 0 to 1 for three products. Then, the class of the product with the highest spatial

438 correspondence was chosen for each pixel.

439 Furthermore, we compared the mapping performance of DSET with Random Forest (RF) which is

440 considered one of the most popular algorithms for land cover mapping. In the land cover classification

441 using the FR classifier, all available Level-2 Tier 1surface reflectance (SR) data of Landsat 8 OLI

442 (Operational Land Imager) sensors from the year 2015 and two adjacent years on GEE was employed.

443 All Landsat images have been atmospherically corrected. The following six bands were used as input

444 features: blue, green, red, NIR, SWIR1, and SWIR2. To improve the mapping performance, several

445 important spectral indices, including DNVI, NDWI, and NDBI were also used as auxiliary data to the

446 RF classifier. The RF classifier was trained on 80% of the global point-based samples since those samples

447 were of high quality after manual visual interpretation of high-resolution images. As the global land cover

448 mapping based on the RF classifier is a tough task, we randomly selected a total of 300 grids with the

449 size of 4° (Figure S2) and applied corresponding local RF classifiers to these grids. Then, the mapping

450 results were validated by the remaining 20% of the point-based samples.

451 **4. Results and discussion**

452 **4.1 Mapping result of the GLC-2015 map**

453 Using a multi-source product fusion method based on the DSET, we generated an improved 30m global

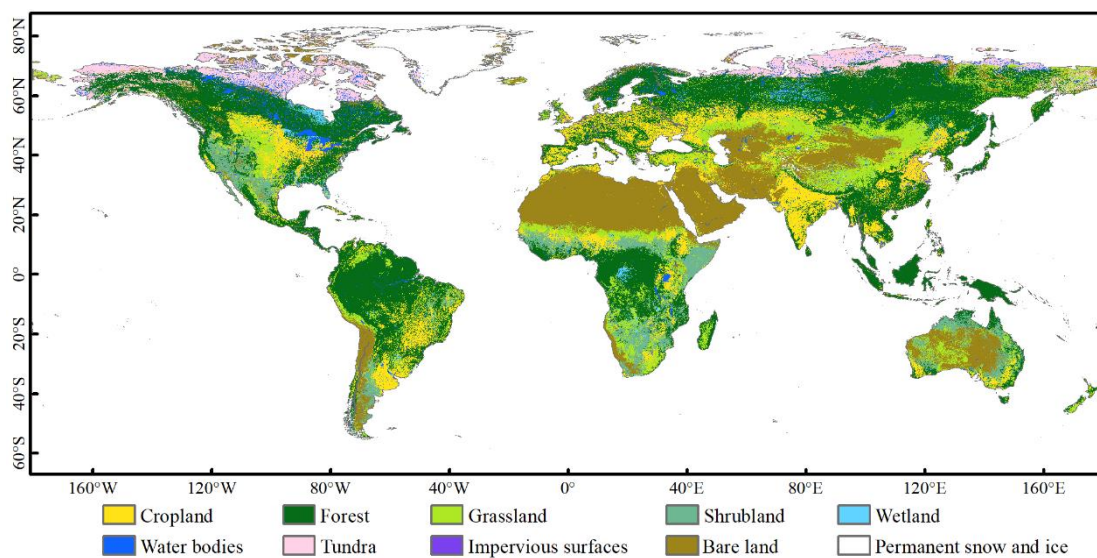
454 land cover map in 2015 (GLC-2015). Figure 5 illustrates the GLC-2015 map. The GLC-2015 map can

455 accurately describe the spatial distribution of various LC classes. For example, cropland areas are mostly

456 located in Central America, the region from the Hungarian plain to the Siberian plain, the eastern and

457 southern parts of China, and the most of India. In addition, forest, which is one of the easily

458 distinguishable classes from the map, is concentrated in the eastern part of North America, the Amazon

459 basin of South America, the northern part of Eurasia and the equatorial region of Africa.

460

461 Figure 5. Global land cover map in 2015 with 30 m resolution (GLC-2015).

462 **4.2 Accuracy assessment of the GLC-2015 map**

463 **4.2.1 Accuracy assessment with the global point-based samples**

464 The accuracy of the GLC-2015 map was first tested via the global point-based samples, and the results

465 of assessment are listed in Table 3. The GLC-2015 map achieved an OA of 79.5% and kappa coefficient

466 of 0.757 at the global scale, demonstrating the good performance of our map. Among all the LC classes,

467 permanent snow and ice possessed the best mapping performance, with PA and UA achieving 89.1% and

468 93.7%. The accuracy of water bodies, forest and impervious surfaces was also high, where PA and UA

469 exceeded 80.0%. Grassland, shrubland, and wetland had relatively low accuracy, with PA below 75.0%.

470 Among them, grassland and shrubland were mainly confused with forest, which might be because these

471 classes are both vegetation, thus causing difficulty in recognition by spectral information. Due to the

472 complex spectral characteristics, wetland is often mixed with vegetation (Ludwig et al., 2019).

473 **Table 3. The confusion matrix for the GLC-2015 map based on the global point-based samples.**

	Cropland	Forest	Grassland	Shrubland	Wetland	Water bodies	Tundra	Impervious surfaces	Bare land	Permanent snow and ice	Total	PA
Cropland	3623	387	356	61	27	48	2	71	81	0	4656	0.778
Forest	155	8813	186	141	232	16	43	43	53	3	9685	0.910
Grassland	10	337	1920	19	24	13	47	36	184	9	2599	0.739
Shrubland	155	438	656	1469	39	29	70	78	442	4	3380	0.435
Wetland	47	287	82	14	1067	64	22	18	110	4	1715	0.622
Water bodies	27	90	15	1	73	1936	17	10	44	3	2216	0.874
Tundra	1	242	119	6	29	19	1411	2	269	17	2115	0.667
Impervious surfaces	74	41	11	3	8	11	1	1295	45	0	1489	0.870
Bare land	36	59	237	32	44	91	55	60	4909	38	5561	0.883
Permanent snow and ice	0	11	8	0	4	18	13	1	86	1154	1295	0.891
Total	4128	10705	3590	1746	1547	2245	1681	1614	6223	1232	34711	
UA	0.878	0.823	0.535	0.841	0.690	0.862	0.839	0.802	0.789	0.937		
OA						0.795						
Kappa						0.757						

474 The regional accuracies are presented in Figure 6. The OA of the GLC-2015 ranged from 66.4% to

475 93.4%, and kappa coefficient from 0.552 to 0.813. From the perspective of OA, Water regions lead,

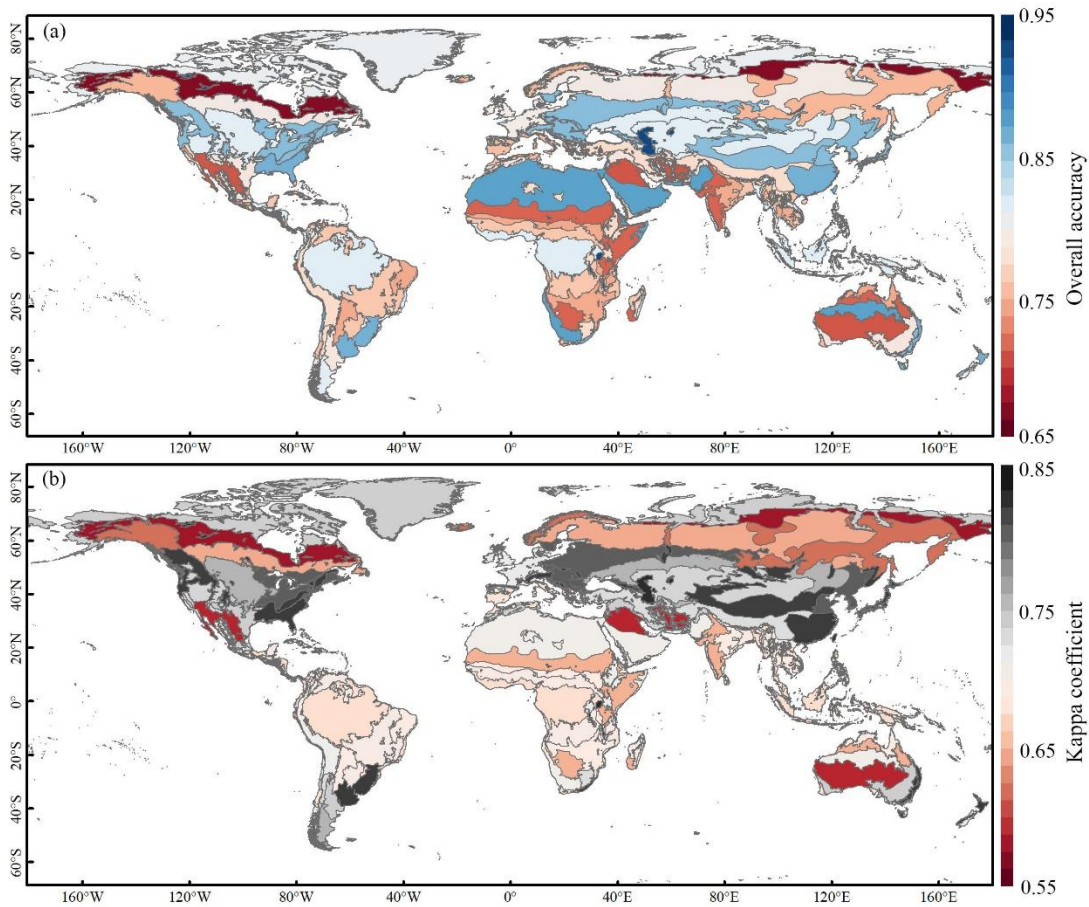
476 followed by Tropical desert, Temperate continental forest, and Polar. These are areas with homogeneous

477 land cover and have low difficulty in mapping. Boreal tundra woodland, Tropical dry forest, Tropical

478 shrubland, and Subtropical desert are the regions with low OA. The first one may be related to the high

479 latitudes. The followed two may be because they belong to areas with complicated and mixed LC classes

480 which is not easily classified. The last one may be the consequence of sparse vegetation in desert areas.

481   For the kappa coefficient, the ranking was similar with those for OA.

482
483   **Figure 6. Regional accuracy of the GLC-2015 map according to ecoregions. (a)overall accuracy, (b) kappa**
484   **coefficient. The ecoregion boundaries are obtained from the Food and Agriculture Organization of the United**
485   **Nations (FAO).**

486   **4.2.2 Accuracy assessment with the global patch-based samples**

487   The accuracy assessment of the GLC-2015 map was also conducted with the global patch-based samples.

488   Table 4 summarizes the results for accuracy assessment of each LC class in the GLC-2015 map. From

489   the assessment results, it can be found that the OA of the GLC-2015 map reached 83.6%, which was

490   higher than 79.5% tested with the global point-based samples. The kappa coefficient of the GLC-2015

491   map was 0.566, which was 0.191 lower than the result calculated with the global point-based samples.

492   In both accuracy assessment results based on two different validation data sets, water bodies, forest, and

493   permanent snow and ice were validated to have high accuracy, and grassland, shrubland, and wetland

494   were validated to have low accuracy. Nevertheless, the ranking of accuracy for each LC class had a slight

495   difference. For example, in assessment based on the global point-based samples, impervious surfaces

496   and permanent snow and ice ranked higher than that based on the global patch-based samples. This may

497   be because a LC map can easily show where one LC class is distributed but hardly describe its actual

498   shape. In addition to the accuracy assessment on a pixel scale, validation on a patch scale is equally

499   important because it can reflect the shape consistency between the GLC-2015 map and the actual

500   landscape, even if the size of global patch-based samples is relatively small. Overall, no matter from the

501   respective of the global point-based samples or the global patch-based samples, the mapping accuracies

502   of the GLC-2015 map are satisfactory.

503   **Table 4. Mapping accuracy via the global patch-based samples for the GLC-2015 map**

	Cropland	Forest	Grassland	shrubland	Wetland	Water bodies	Tundra	Impervious surfaces	Bare land	Permanent snow and ice
PA	0.887	0.895	0.629	0.589	0.301	0.939	0.701	0.757	0.682	0.825
UA	0.916	0.844	0.617	0.714	0.511	0.917	0.872	0.713	0.599	0.767
OA							0.836			
Kappa							0.566			

504   **4.3 Inter-comparison with other GLC products**

505   **4.3.1 Inter-comparison based on the global point-based samples**

506   Based on the global point-based samples, the inter-comparison of the GLC-2015 map with the

507   GlobeLand30, FROM_GLC, and GLC_FCS30 were conducted. Since the three products used different

508   classification systems, LC classes were transformed to the classification system we adopted in this paper

509   to achieve consistent accuracy assessment. The accuracy assessment results for all GLC maps are listed

510   in Table 5. It can be found that the GLC-2015 map achieved the highest OA of 79.5% compared with

511   GlobeLand30 of 65.3%, FROM_GLC of 61.7%, and GLC_FCS30 of 65.5%, respectively. The accuracy

512   gap between the GLC-2015 map and other existing ones was 14.0%-17.8%. Also, the GLC-2015 map

513   possessed a better kappa coefficient than other products. For each LC class, the GLC-2015 map

514   outperformed the other three maps in terms of PA except for tundra. For cropland, grassland, shrubland,

515   wetland, and tundra, the GLC-2015 map also exhibited better performance for UA than the GlobeLand30,

516   FROM_GLC and GLC_FCS30. Overall, for the PA or UA, the GLC-2015 map ranked first or second in

517   nearly all LC classes, which demonstrated that the GLC-2015 map had smaller omission and commission

518   errors against the other three products.

519   **Table 5. Mapping accuracy of the GLC products with the global point-based samples.**

		Cropland	Forest	Grassland	Shrubland	Wetland	Water bodies	Tundra	Impervious surfaces	Bare land	Permanent snow and ice	OA (Kappa coefficient)
GLC-2015	PA	0.778	0.910	0.739	0.435	0.622	0.874	0.667	0.870	0.883	0.891	0.795
	UA	0.878	0.823	0.535	0.841	0.690	0.862	0.839	0.802	0.789	0.937	(0.757)
Globeland30	PA	0.752	0.719	0.713	0.245	0.540	0.680	0.769	0.688	0.609	0.821	0.653
	UA	0.786	0.818	0.255	0.428	0.573	0.869	0.577	0.809	0.868	0.905	(0.598)
FROM_GLC	PA	0.389	0.694	0.707	0.411	0.307	0.607	0.712	0.732	0.731	0.881	0.617
	UA	0.671	0.859	0.278	0.422	0.289	0.742	0.686	0.661	0.761	0.773	(0.558)
GLC_FCS30	PA	0.757	0.775	0.452	0.399	0.455	0.604	0.228	0.777	0.809	0.726	0.655
	UA	0.616	0.816	0.384	0.405	0.515	0.808	0.688	0.774	0.645	0.947	(0.591)

520   Further quantitative accuracy assessments of different GLC products were performed in 4° × 4°

521   grids using the global point-based samples, and box plots were produced for each product for all grids

522   within different ecoregions, as shown in Figure 7. It can be found that the GLC-2015 map outperformed

523   other existing products with the best OA and kappa coefficient across different ecoregions. Also, the

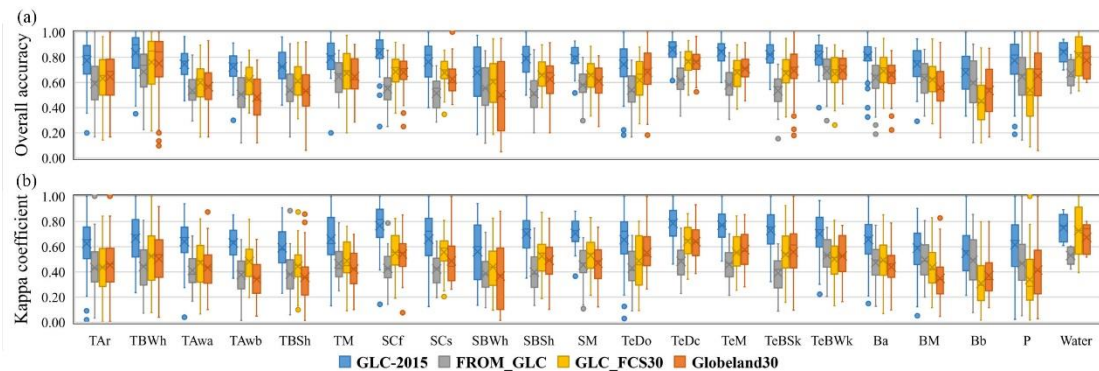
524   mean overall accuracy of the GLC-2015 map exceeded 65.0% in all ecoregions, showing the high quality

525   of our mapping results. It is worth noting that the GLC-2015 map showed shorter boxes except in

526   Subtropical dry forest and Subtropical desert, which means the GLC-2015 map had relatively small

527   fluctuation than other ones. In Subtropical desert, Tropical dry forest, and Boreal tundra woodland, the

528   OA and kappa coefficient of the four products were relatively low. However, the GLC-2015 map

529   exceeded the highest of others and greatly improved the mean OA in these regions.

530

531 **Figure 7. The box-plot of the accuracy for twenty-one ecoregion zones. (a) overall accuracy, (b)kappa**

532 **coefficient. Ecoregion abbreviation and corresponding ecoregion is described in Table S4.**

533 **4.3.2 Inter-comparison based on the global patch-based samples**

534 Although the global point-based samples are adequate and even across the globe, the distribution of

535 points in each 4° × 4° geographical grid is too sparse to reflect the actual spatial pattern of the landscape.

536 Focusing on LC pattern at the local scale, we also used the global patch-based samples which can provide

537 spatial context information to conduct the accuracy assessment of the GLC-2015 map and compare

538 difference GLC products. Table 6 lists the accuracies of the GLC-2015 map and the other three GLC

539 products. Obviously, the GLC-2015 map achieved the best OA and kappa coefficient among these four

540 GLC maps. The overall accuracy gap between the GLC-2015 product and others was 5.9%-24.5%, which

541 presented a more significant variation compared with the result based on the global point-based samples.

542 In terms of PA and UA, the GLC-2015 map was higher than the other three ones in most LC classes.

543 Specifically, all the products had lower accuracy for grassland, shrubland, and wetland, similar to that in

544 the accuracy assessment based on the global point-based samples. It is evident that the FROM_GLC had

545 the lowest mapping accuracy for grassland, shrubland, and wetland, implying that the classification

546 method of FROM_GLC is not robust for these three LC classes.

547 **Table 6. Mapping accuracy of the GLC products with the global patch-based samples**

		Cropland	Forest	Grassland	Shrubland	Wetland	Water bodies	Tundra	Impervious surfaces	Bare land	Permanent snow and ice	OA
GLC-2015	PA	0.887	0.895	0.629	0.589	0.301	0.939	0.701	0.757	0.682	0.825	0.836
	UA	0.916	0.844	0.617	0.714	0.511	0.917	0.872	0.713	0.599	0.767	(0.566)
Globeland30	PA	0.896	0.698	0.765	0.539	0.455	0.824	0.752	0.643	0.492	0.831	0.777
	UA	0.891	0.906	0.444	0.527	0.157	0.893	0.500	0.703	0.829	0.705	(0.437)
FROM_GLC	PA	0.485	0.714	0.640	0.254	0.032	0.904	0.760	0.506	0.681	0.501	0.591
	UA	0.872	0.809	0.193	0.139	0.186	0.884	0.696	0.808	0.496	0.703	(0.360)
GLC_FCS30	PA	0.865	0.779	0.398	0.565	0.363	0.869	0.051	0.648	0.658	0.742	0.748
	UA	0.857	0.832	0.509	0.330	0.132	0.942	0.573	0.643	0.462	0.752	(0.418)

548 Accuracy assessment was calculated in each patch-based sample, and box plots were produced for

549 each GLC product at the continental scale, as shown in Figure 8. The GLC-2015 map showed a robust

550 performance in each continent, with the highest OA and kappa coefficient among all the maps. Also, in

551 all continents, the GLC-2015 map had the shortest boxes in terms of OA, which denoted that it had a

552 more minor variation in accuracy at the continental scale. Among four products, the GLC_FCS30 and

553 Globeland30 achieved similar accuracies in most continents. Obviously, the FROM_GLC showed lowest

554 accuracy across different continents, especially in Oceania, where the OA of most patch-based samples

555 was below 40.0%, namely most of the pixels in Oceania were incorrectly classified. We further compared

556 mapping accuracies for each LC class in different continents (Figures S3-S4). Since tundra and

557 permanent snow and ice are rare and only existent in certain regions, they were not included in the

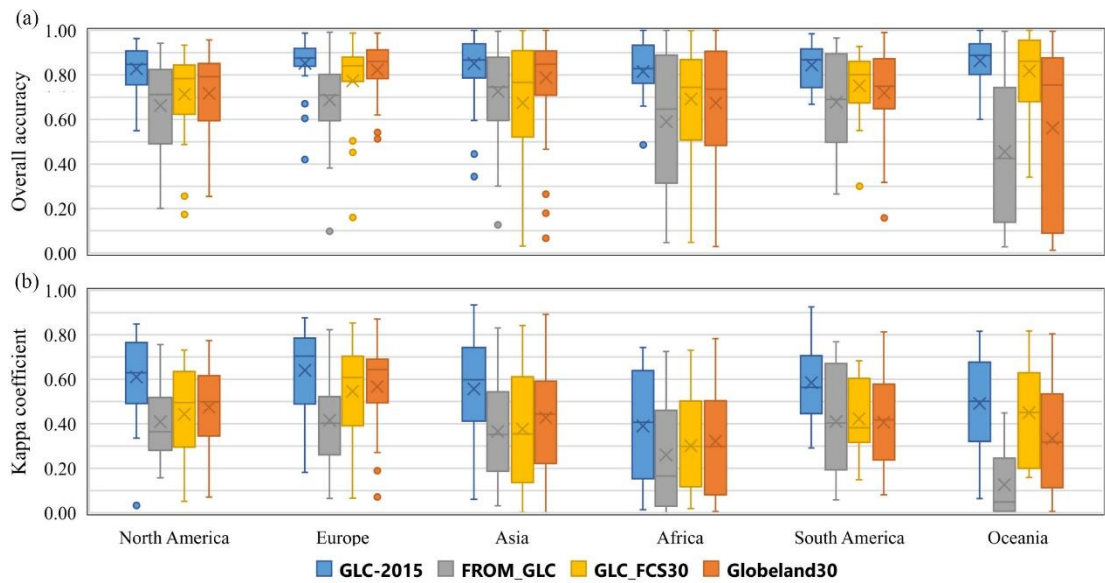
558 comparison. As for PA across different continents, the GLC-2015 map outperformed other maps in forest,

559 water bodies, and bare land. As for UA across different continents, the GLC-2015 map outperformed

560 other maps in cropland, grassland, shrubland and wetland, and achieved similar accuracies with the

561 GLC_FCS30 and Globeland30 in forest. Overall, the GLC-2015 map outperformed others regarding

562 mapping accuracy at continental scale. In addition, all GLC products showed significant variation and

563 low mean accuracy in grassland, shrubland, and wetland over most continents.

564
565 **Figure 8. The box-plot of the accuracy for different continents. (a) overall accuracy, (b)kappa coefficient.**

566 Furthermore, to compare the OA of the GLC-2015 map with other GLC products, scatter plots were

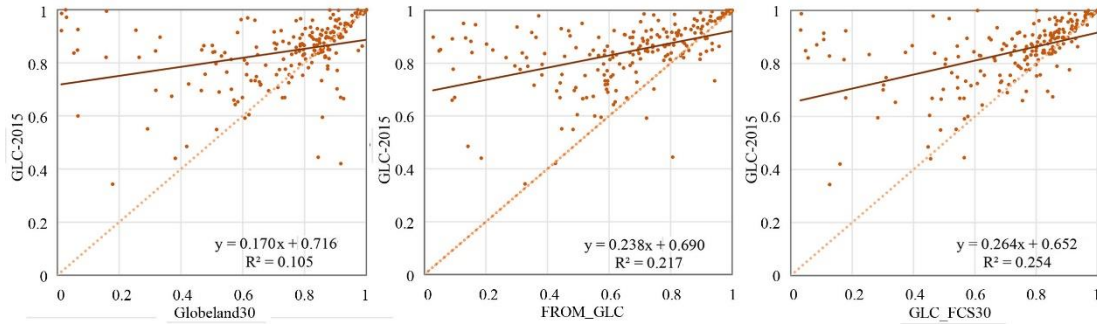
567 used to describe the relationship between the overall accuracy of the GLC-2015 map and one other

568 product in each patch-based sample, as displayed in Figure 9. Most of the points were above the 1:1 line,

569 implying that the GLC-2015 map surpassed other GLC products in terms of OA. The distribution of

570 points was more dispersed from the 1:1 line in the plot of the GLC-2015 map against FROM_GLC

571 compared to other plots. It indicated that these two products had a more significant difference, which

572 was also proved in Table 6.

573

574 **Figure 9. Scatter plots between the GLC-2015 map and other products obtained using the global patch-based**

575 **samples.**

576 **4.3.3 Areal comparison for individual classes**

577 To assess the similarities and discrepancies between the GLC-2015 and other GLC products, we

578 compared the area of various LC classes at multiple scales, including global, continental, national, and

579 ecoregional scales.

580 The areal comparison for various classes of different GLC products over the globe is shown in

581 Figure 10. Generally, the areas of water bodies and permanent snow and ice of four GLC products were

582 very similar, which may be related to the similar LC definitions. In contrast, the areas of cropland, forest,

583 grassland, and shrubland in GLC-2015 differed significantly from those in other GLC products. The area

584 of forest in GLC-2015 is much higher than other products. This may be because FROM_GLC and

585 GLC_FCS30 defined forest with tree cover over 15%, while GLC-2015 used a threshold of over 10%.

586 The cropland areas in GLC-2015 and Globeland30 were close, higher than FROM_GLC but lower than

587 GLC_FCS30. Moreover, the FROM_GLC underestimated the cropland area as it had a low producer's

588 accuracy for cropland (see Table 5), which was also demonstrated in previous researches (Liu and Xu,

589 2021; Zhang et al., 2021). FROM_GLC and Globeland30 shared similar grassland areas since a similar

590 accuracy for grassland was found in these two products (see Table 5). However, the FROM_GLC and

591 Globeland30 significantly overestimated grassland extent, with much bare land misclassified as

592 grassland (Hu et al., 2014). The GLC_FCS30 showed the smallest area for grassland, which might be

593 related to its higher threshold in vegetation cover for grassland. For shrubland, the area difference

594 between GLC-2015 and Globeland30 was minimal, and the areas in FROM_GLC and GLC_FCS30 were

595 similar. Furthermore, the wetland area in FROM_GLC was the lowest among all the products, with a

596 total area of 0.168 million km^2. In contrast, the Globeland30 and GLC_FCS30 exhibited greater wetland

597 extent than GLC-2015 since these two products classified non-wetlands sensitive to water as wetlands

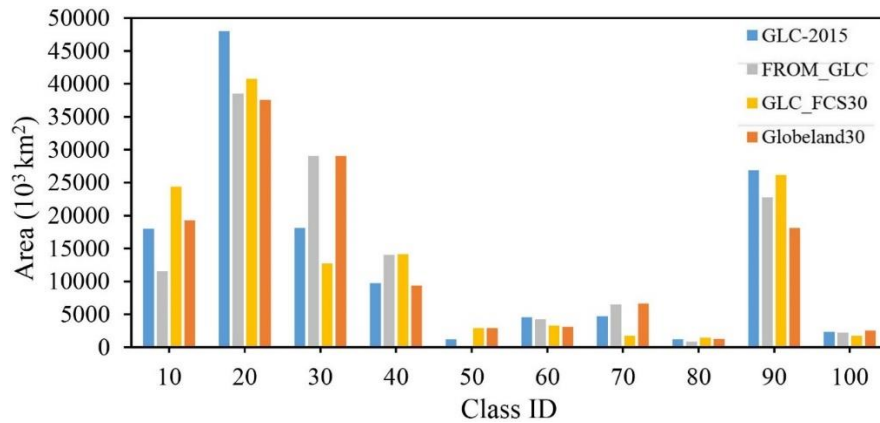
598   (Zhang et al., 2023). In particular, the tundra area in GLC_FCS30 was much smaller than other products.

599   This is mainly because only lichens/mosses in the original classification system of GLC_FCS30 was

600   converted into tundra in the classification system we used, which leads to the omission of tundra. The

601   areas of impervious surfaces in GLC-2015, Globeland30, and GLC_FCS30 were very close and higher

602   than FROM_GLC. For bare land, there was large difference between Globeland30 and other products,

603   while the area in GLC-2015 and GLC_FCS30 was very close.

604

605   **Figure 10. Areal comparison of various land cover classes among GLC products at the global scale. Class IDs**

606   **10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 denote cropland, forest, grassland, shrubland, wetland, water bodies,**

607   **tundra, impervious surfaces, bare land, and permanent snow and sea ice, respectively.**

608   The area similarity and difference for various classes of different GLC products were also compared

609   over six continents, the top 40 countries ranked by area, and 21 ecoregions (Figures S5-S7). Overall, the

610   four products showed a similar distribution trend of different classes. For most LC classes, the continental,

611   national, and ecoregional rankings of four products agreed with their ranking at the global scale. Whereas,

612   for grassland and shrubland, the area ranking of four products varied at three different regional scales.

613   **4.3.4 Visual inter-comparison for individual classes**

614   The visual comparison of cropland in GLC-2015, Globeland30, FROM_GLC, GLC_FCS30, GSFAD30

615   (Xiong et al., 2017; Teluguntla et al., 2018), and other national-scale maps was conducted in three local

616   regions (Figure S8). In the Egyptian agricultural area, GLC-2015, FROM_GLC, and GLC-FCS30 shared

617   similar delineation of the cropland and had a good representation of cropland with fine spatial details.

618   Since the date time of the Google Earth image is 2015, Globeland30 missed the newly cultivated cropland.

619   GFASD30 had the largest cropland area among five products but misclassified bare land as cropland. In

620   the agricultural area of Southeastern China, GLC-2015 had an agreement with GFSAD30 and CLCD.

621   Globeland30 and GLC_FCS30 overestimated the area of cropland. As for FROM_GLC, it failed to depict

622   the spatial distribution of cropland and had many omissions. In cropland-dominated areas of the United

623   States, FROM_GLC significantly underestimated the extent of cropland. The other five products

624   exhibited a similar delineation of cropland, but there were little differences in some small areas. For

625   example, Globeland30 misclassified some grassland into cropland, and NLCD 2016 had a good ability

626   to distinguish the farm rack.

627   We also compared the performance in the forest of different products in three forest-prevalent

628   regions of Congo, China, and the United States (Figure S 9). Overall, GLC-2015 and Globeland30

629   showed accurate delineation in three regions. FROM_GLC also had good performance for the forest in

630   Congo and USA but overestimated the forest in China, mislabeling shrubland and grassland as forest.

631   Furthermore, GFC tended to miss sparse trees in China, and GLC_FCS30 underestimated the extent of

632   forest in both three regions. As for national-scale products, CLCD and NLCD 2016 had a good ability to

633   identify the details of forest, while CLUD dramatically missed both dense and sparse woodlands.

634   Furthermore, to compare the performance in the wetland of GLC-2015 with other global and

635   national-scale products, three wetland regions in South-central Canada, coastal America, and Sundarbans

636   were selected. It can be found that GLC-2015 and Globeland30 had similar representation and performed

637   well in identifying the wetland over three regions (Figure S10). Unexpectedly, FROM_GLC performed

638   poorly in each region, with almost no wetlands captured. GLC_FCS30 also showed unstable quality in

639   three regions. For example, it highly underestimated the wetland area in coastal America and completely

640   mislabeled the mangroves as cropland in Sundarbans. NLCD 2016 and GMW accurately demonstrated

641   the spatial pattern of the wetland, while the CA_wetlands map underestimated the wetland extent because

642   it defined wetlands by wetland frequency of no less than 80% from 2000 to 2016 (Wulder et al., 2018).

643   To understand the spatial distribution of impervious surfaces in different products, a comparison of

644   mapping results for three megacities, including Tokyo, Shanghai, and New York, was shown in Figure

645   S11. In Tokyo, a high consistency was found between GLC-2015, FROM_GLC, and GAUD, and both

646   successfully captured the impervious surfaces in peri-urban areas. GLC_FCS30 showed the largest area

647   for impervious surfaces because it misclassified many croplands into impervious surfaces. In Shanghai,

648   GLC_FCS30 underestimated the central city, and CLUD lost the details of impervious surfaces because

649   it was developed using the visual interpretation method. Other products generally had the similar

650   representation and accurately demonstrated the spatial distribution of the city. For New York, the

651   FROM_GLC, GLC_FCS30, and GAUD agreed well with GLC-2015, while Globeland30 and NLCD

652   2016 had high impervious areas than others.

4.3.5 Visual inter-comparison at the local scale

654   Except for visual comparison for individual classes among various products, we selected six typical

655   geographical tiles covering six continents and different landscape environments to further present the

656   mapping performance of the GLC-2015 map, Globeland30, FROM_GLC, and GLC_FCS30, as shown

657   in Figure 11. Overall, from a local point of view, the GLC-2015 map tended to be more diverse in LC

658   classes and had better identification performance in various classes. In flattened cropland areas (Figure

659   11a and Figure 11b), the GLC-2015 map revealed diverse LC classes and accurately distinguished

660   impervious surfaces; however, the Globeland30 exaggerated the extent of impervious surfaces, and the

661   FROM_GLC failed to delineate impervious surfaces with small size. In addition, the FROM_GLC

662   misclassified some cropland pixels as grassland (Figure 11a) and had an abnormal "stamp" (Figure 11b).

663   As for mountain areas (Figure 11c and Figure 11d), the GLC-2015 map uncovered the spatial pattern of

664   natural and planted forest, cropland, and grassland. There were large confusions between cropland and

665   grassland in the results of the FROM_GLC and GLC_FCS30, and some impervious surfaces and

666   cropland areas were wrongly labeled as bare land by the FROM_GLC. The areas (Figure 11c), which

667   were classified as forest, were misidentified as cropland and grassland in three other products. For the

668   rainforest areas where a large number of trees were reclaimed for cropland (Figure 11e), the GLC-2015

669   map, Globeland30, and GLC_FCS30 had similarities in cropland areas; but the FROM_GLC recognized

670   some reclaimed areas as grassland. Additionally, the GLC-2015 map accurately presented the spatial

671   distribution of impervious surfaces while other products had omission or commission errors. In the

672   cropland-dominated areas (Figure 11f), the GLC-2015 map and Globeland30 showed a higher agreement,

673   and both of them mapped the undulating areas as grassland. Unlike the aforementioned two products, the

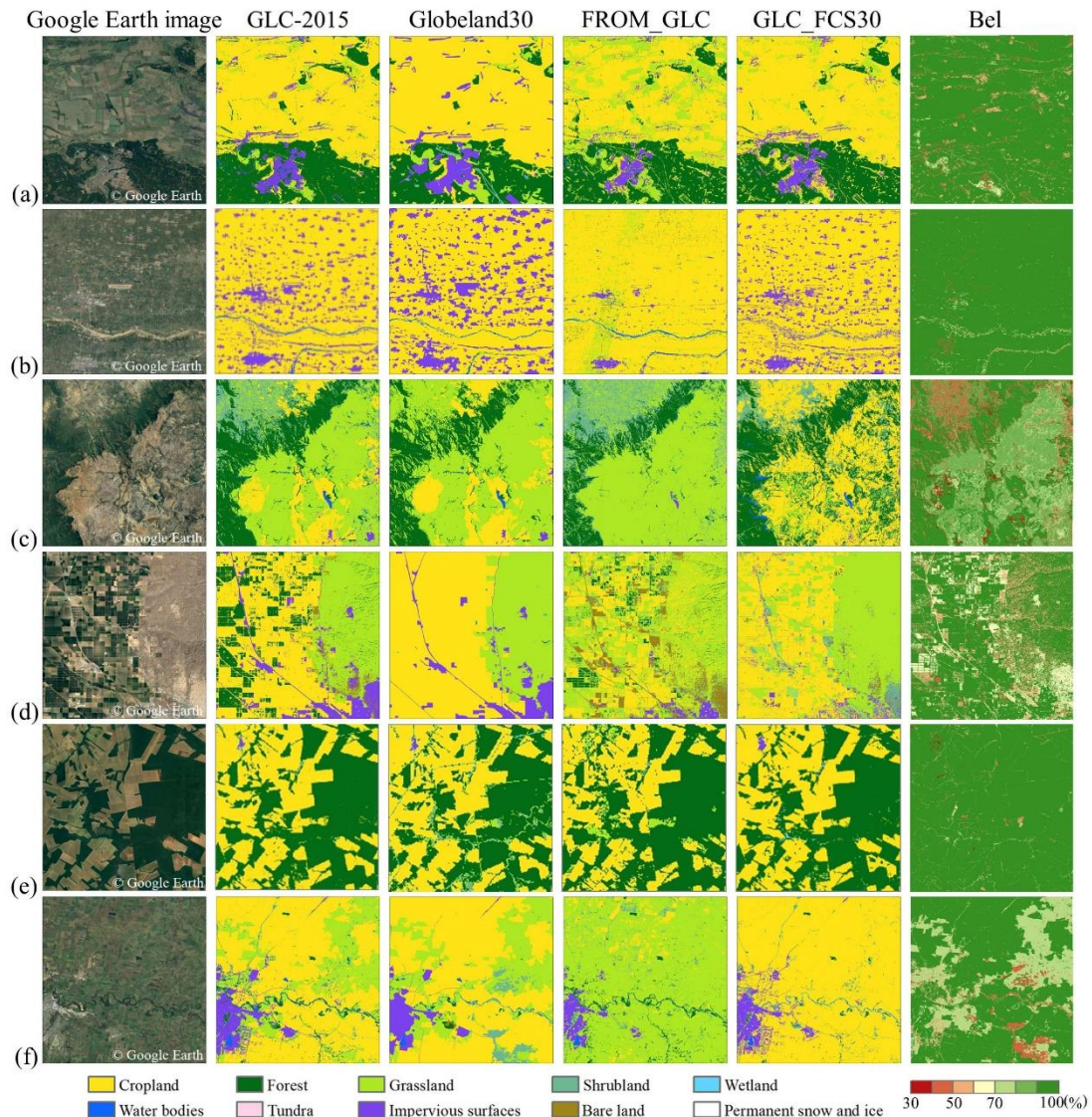
674   FROM_GLC misclassified large tracts of croplands as grasslands, and the GLC_FCS30 did not capture

675   the grassland in undulating areas. Figure 11 also shows the belief measure of the fused result in different

676   geographical tiles. Although it does not directly evaluate the mapping accuracy, it serves as a degree of

677   support for the hypothesis of an accepted LC class being true, it can still reflect the quality of the GLC-

678   2015 map. Overall, Bel of the GLC-2015 map exceeded 80% in most areas of each tile, demonstrating

679   the credibility and high quality of our mapping result.

683 **4.4 Improvement of the GLC-2015 map compared to other GLC products**

684   The spatial distribution of inconsistency between three GLC products at the global scale is illustrated in

685   Figure 12. From the inconsistency map, we found that areas of low inconsistency mainly corresponded

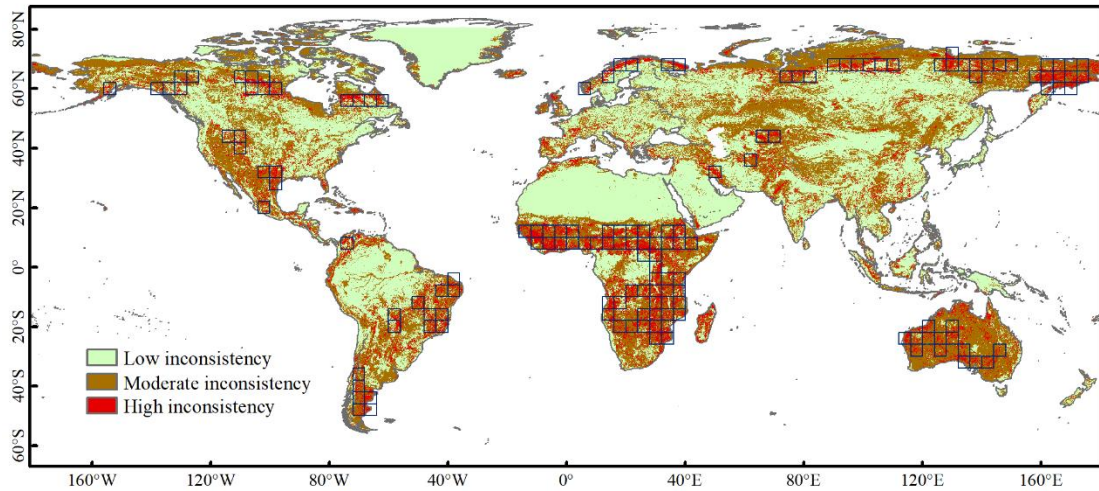
686   to homogeneous regions with simple LC classes. For example, the northern part of Africa was mainly

687   classified as bare land, the northern part of South America was mainly classified as forest, and the

688   Greenland was classified as permanent snow and ice. On the contrary, areas of high inconsistency were

689   located in regions with complicated LC classes, especially in mixed vegetation regions or sparse

690   vegetation regions, such as northern Asia, South Africa, Sahel region, Australia, northern and southern

691   North America, and eastern and southern South America.

693   **Figure 12. Distribution of inconsistency between the Globeland30, FROM_GLC, and GLC_FCS30.The blue**
694   **rectangles are high-inconsistency grids that the area of pixels with value equal to 1 account for more than 20%**
695   **of the total area.**

696   Based on the global point-based samples, we assessed the accuracies of the GLC-2015 map,

697   Globeland30, FROM_GLC, and GLC_FCS30, in the aforementioned areas of low inconsistency,

698   moderate inconsistency, and high inconsistency, as shown in Table 7. Overall, the GLC-2015 map had

699   the highest accuracies against the other three ones in three areas. For each product, areas of low

700   inconsistency obtained the highest accuracies, followed by areas of moderate inconsistency and then high

701   inconsistency, which demonstrated that inconsistency of the existing products could indicate the quality

702   of maps. In areas of low inconsistency, the overall accuracy gap between the GLC-2015 map and

703   previous ones was as small as 0.1%-0.6%. However, for areas of moderate and high inconsistency, the

704   comparison accuracy gap expanded to 19.3%-28.0% and 27.5%-29.7%, respectively. It proved the

705   outperformance of the GLC-2015 map over the other three products in the areas of high identification

706   difficulty.

707   **Table 7. Accuracy assessments of the GLC products in three areas.**

	GLC-2015		Globeland30		FROM_GLC		GLC_FCS30	
	OA	Kappa	OA	Kappa	OA	Kappa	OA	Kappa
Areas of low inconsistency	0.951	0.938	0.945	0.929	0.950	0.936	0.951	0.937
Areas of moderate inconsistency	0.760	0.723	0.561	0.498	0.480	0.411	0.567	0.495
Areas of high inconsistency	0.567	0.498	0.292	0.204	0.286	0.198	0.270	0.160

708   We further provided a comparative analysis of three previous GLC products and the GLC-2015 map

709   in areas of high inconsistency. We calculated the area of pixels with a value equal to 1 in 4° × 4° grids.

710   The grids that the area of pixels with a value equal to 1 account for more than 20% of the total area was

711   selected as grids of high inconsistency. Finally, a total number of 147 grids were selected (Figure 13). To

712   compare the accuracy of the GLC-2015 map and other ones, we utilized scatter plots to represent the

713   relationship between the overall accuracy of one previous product and the GLC-2015 map in each grid

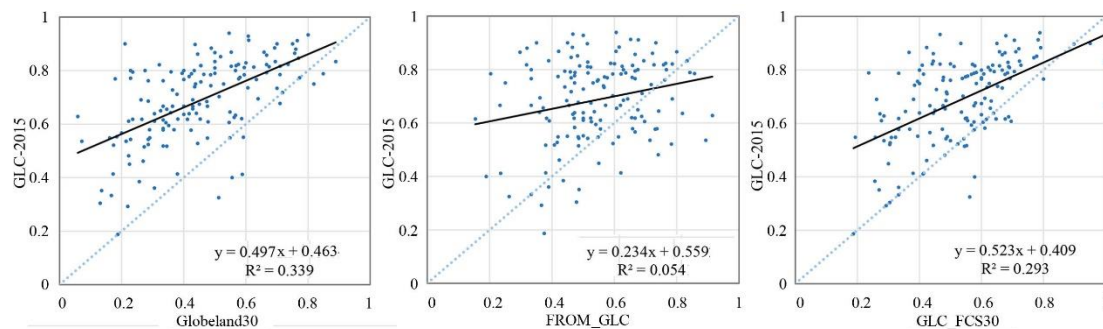
714   of high inconsistency based on the global point-based samples (Figure 13). Most of the points were above

715   the 1:1line, namely the values of y-axes corresponding to those points were larger than the values of x-

716   axes, which demonstrated that the GLC-2015 map performed better than other GLC products in most

717   grids of high inconsistency. It can be found that the fitting line in each scatter plot had the intercept

718   exceeding 0.40, the slope less than 0.55, and the R^2 less than 0.35, showing that the GLC-2015 map had

719   a large difference with other ones.

720

721   **Figure 13. Overall accuracy relationship between the GLC-2015 map and other products in grids of high**

722   **inconsistency.**

723   To intuitively compare the mapping result of the GLC-2015 map and three existing ones in areas of

724   high inconsistency, we focused on visual inspection in various areas based on four 5 km×5km patch-

725   based samples and conducted accuracy statistics, as shown in Figure 14. In the detailed display, it is

726   apparent that three previous products had a large difference in four areas. As can be seen from the four

727   visual cases, the typical confusions between LC classes in areas of high inconsistency were as follows:

728   (1) shrubland was easily misclassified as forest and grassland; (2) cropland, grassland, and shrubland

729   were heavily confused with each other; (3) bare land was likely to be mixed with shrubland and grassland.

730   Overall, the GLC-2015 map surpassed other products in the local accuracy assessment. In Western

731   Australian mulga shrublands (Figure 14a), the GLC-2015 map and GLC_FCS30 showed similar spatial

732   distribution and shape of bare land and forest, which was consistent with the real landscape. While the

733   Globeland30 classified bare land as grassland and the FROM_GLC under-classified bare land. As for

734   Zambezian and mopane woodlands (Figure 14b), the GLC-2015 map performed best with OA reaching

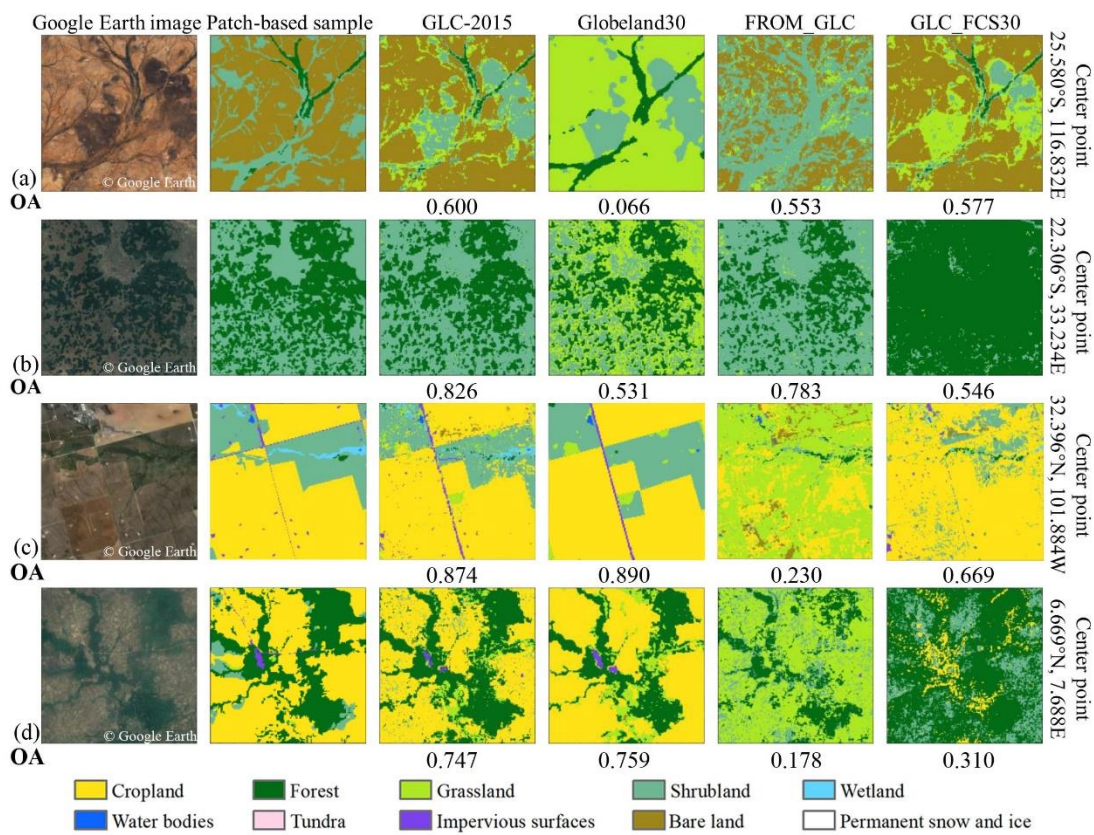
735   82.6%, followed by the FROM_GLC. In contrast, other products mixed shrubland with forest or

736   grassland. In agricultural land of Western United States (Figure 14c), the GLC-2015 and Globeland30

737   exhibited similar mapping results with the ground truth while the FROM_GLC had large difference with

738   other products. When it comes to Guinean forest-savanna mosaic (Figure 14d), the GLC-2015 map and

739   Globeland30 showed high spatial consistency, and both had accurate classification profile for cropland,

740   forest, and impervious surfaces, while other products misidentified cropland as other LC classes.

741

742   **Figure 14. Visual comparison between the GLC-2015 map and three other products based on 5km × 5km**

743   **patch-based samples and Google Earth images for four areas of high inconsistency (a-d). The OA for each**

744   **product was calculated by the corresponding patch-based sample.**

745   **4.5 Comparison between DSET and other methods**

746   **4.5.1 Inter-comparison with other data fusion methods**

747   The accuracy assessments on GLC-2015 obtained by DSET and global mapping results from two other

748   data fusion methods were conducted based on two global validation sample sets. The confusion matrixes

749 with the global point-based samples are shown in Table S5 and S6. The OA of the global land cover

750 classification obtained by the MV and SC was 72.1% and 71.8%, respectively. As shown in Table 3, the

751 OA of the GLC-2015 map obtained by the DSET method was 79.5%, which had an improvement of 7.4%

752 and 7.7% compared to mapping results from the MV and SC. In addition, the GLC-2015 map obtained

753 higher PA and UA for most LC classes.

754   When evaluating GLC maps obtained by different data fusion approaches using the global patch-

755 based samples, the DSET method obtained the highest OA of 83.6% and kappa coefficient of 0.566,

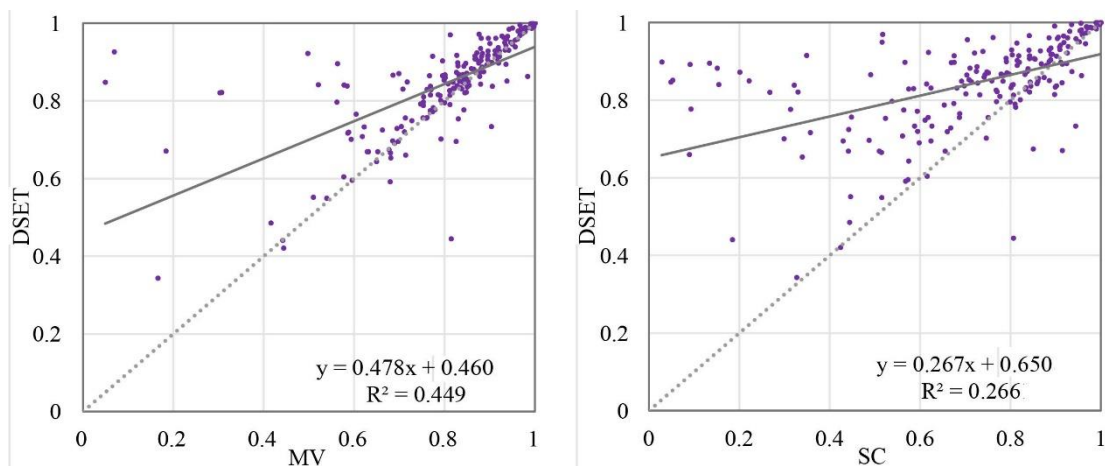
756 compared with 80.1% and 0.497 for MV, and 71.8% and 0.391 for SC (Table S7). Here, the DSET method

757 achieved an accuracy improvement of 3.5% and 11.8%. Compared to the two other methods, the DSET

758 improved the accuracy for nearly all the LC classes, especially for grassland, shrubland, and wetland.

759 We also compared the overall accuracy relationship between the DSET and other methods. From the

760 scatter plots (Figure 15), we found that the majority of points were above the 1:1 line, implying DSET

761 had better mapping performance than others in most regions across the globe.

762

763 **Figure 15. Scatter plots between the DSET and other data fusion methods based on the global patch-based**

764 **samples.**

765   Land cover mapping results from the DSET and other methods were also visually illustrated in six

766 tiles with size of the 0.25° covering different continents, as displayed in Figure S12. Despite that mapping

767 results from the DSET and MV depicted similar spatial distribution of LC classes in all tiles except the

768 tile in North America, the DSET more accurately delineated the impervious surfaces of small size which

769 scattered in cropland-dominated (Figure S12a) or arid areas (Figure S12c). Notably, the mapping results

770 from the SC method presented significant differences from that obtained by the DSET and MV. For

771 example, the SC method failed to capture scattered rural residential areas (Figure S12b) and misclassified

772 grassland as cropland (Figure S12d). Overall, the DSET method possessed better recognition

773 performance in various LC classes than the other two methods.

774 **4.5.2 Inter-comparison with the Random Forest**

775 Based on the validation data from 20% of the global point-based samples, we evaluated the quality of

776 the GLC-2015 map obtained by the DSET method and mapping results classified by the RF classifier for

777 a total of 300 grids. The DSET method obtained a mean OA of 80.9% across six continents, while the

778 RF achieved a lower accuracy of 69.9%. From the scatter plots which compared the OA and kappa

779 coefficient between the DSET and RF grid by grid, it was found that the DSET possessed higher accuracy

780 in most grids (Figure S13). Especially, the points were clustered in the upper right corner of the plot

781 (Figure S13a), which indicated that the RF classifier trained with the global point-based samples

782 performed well in those selected grids though it was inferior to the DSET method. Figure S14 shows the

783 OA of the DSET and RF across six continents. We found that the DSET method outperformed RF

784 classifier for each continent. Especially, the mapping results of both two methods presented the lowest

785 accuracy in Oceania. It may be because the selected grids are located in regions with heterogeneous

786 landscape. As for the box plot for the RF classifier, the low hinge exceeded 60.00% in all continents

787 except Oceania, demonstrating the reliability of the RF classifier trained by the global point-based

788 samples. Nevertheless, the performance of the RF classifier was worse than the DSET method. This

789 highlights the feasibility of the DSET method in integrating the existing maps for a better one.

790 **4.6 Advancement and Limitations**

791 To address the problem that current 30m GLC products have great inconsistency in heterogeneous

792 areas and low mapping accuracy for spectral similar LC classes, this study adopted a multi-source

793 product fusion approach based on DSET to create an improved global land cover map (GLC-2015). The

794 results show that the GLC-2015 had good mapping performance with OA reaching 79.5% and 83.6%

795 based on two different validation sets. Compared with those existing products, the GLC-2015 greatly

796 improved the accuracy across the globe, especially in areas of high inconsistency with a significant

797 improvement of 27.5%-29.7%. Compared with other commonly used data fusion methods, the adopted

798 DSET approach provided higher OA and kappa coefficient which showed the benefit of the DEST in

799 integrating various land cover data. No matter from the respective of the global point-based samples or

the global patch-based samples, the GLC-2015 showed relatively low accuracy for grassland, shrubland, and wetland compared to other LC classes. Those LC classes are challenging to map at the global scale duo to their spectral similarity to other classes, ambiguous definitions, or variety with regions. However, compared to other existing 30m GLC products, the GLC-2015 map performed better with the PA and OA ranking first or second for grassland, shrubland, and wetland, which indicated the improvement of the GLC-2015 in poorly-mapped LC classes. It was found that the GLC-2015 map had worse performance in areas with more disagreements (Table 7). However, the GLC-2015 map surpassed other products in the areas with different agree of inconsistency. Moreover, the accuracy gap between the GLC-2015 map and other ones in areas of high inconsistency was larger than that in areas with fewer disagreements, implying that the GLC-2015 map provides a more accurate characterization of land cover in poorly-mapped areas. Although the GLC-2015 map was not capable of avoiding all the wrong mapping results caused by the disagreements from the candidate GLC products, it outperformed the existing products from the aspects of mapping accuracy for the easily misclassified classes and areas with great inconsistency.

Although the GLC-2015 map can evidently improve mapping accuracy in inconsistent areas, there are still some uncertainties. First, we used three multiple-class GLC maps and four single-class GLC maps as the source data for integration. Since those products provided information of land cover at the global scale, classification errors inevitably exist in some specific regions. The multisource product fusion method based on DEST depends highly on the quality of those candidate maps such that the inconsistency between those source maps might lead to incorrect classification. Second, the date time of the GlobeLand30 is different from that of other maps. Because of the five-year time interval, there are changes in land cover, which inevitably distort the fusion results. However, the changed areas are tiny compared to the world's terrestrial area. The uncertainties caused by the LC changes are minor than those from classification errors. In addition, the global point-based samples were used to evaluate the reliability of each product. The accuracy of GlobeLand30 was lower than the other products for areas with LC changes. In this case, the fusion depended more on other maps to avoid the errors caused by LC changes. Third, due to the different LC definitions, uncertainties in classification system conversion are inevitable (Zhang et al., 2017), which might cause problems for the fusion based on the DSET method. However, we conducted a reliability evaluation of the candidate maps to reduce the influence of uncertainties in

classification system conversion on the fusion. The point-based samples used for reliability evaluation were labeled referring to the LC definitions in our classification system so that all the maps were evaluated under the criterion of the classification system we used. By the reliability evaluation, the candidate maps were assessed to have lower accuracy for areas with mismatched information. When integrating all the maps grid by grid, the mismatched information would contribute less to the fusion. Lastly, most candidate LC products used a simple classification system without a level-2 classification system, so they made no contributions to a more detailed classification system when they served as source data for data fusion. Although some maps provided detailed LC classification results, such as the GLC_FCS30 and FROM_GLC for 2015, there might be several challenges in the standardization and uniformity of level-2 classification systems due to the large discrepancies in the definition and criteria. Therefore, the GLC-2015 adopted a simple classification system containing 10 major LC classes. In future work, measures will be taken to meet the expectation of a more detailed classification system for GLC mapping. An improved GLC product with a detailed classification system rather than a simple one-level classification system can be further developed based on the highly applicable and general DSET method whenever more products with diverse LC classes are available. Additionally, a feasible framework for the conversion of different level-2 classification systems into a uniform system should be developed.

5. Data availability

The improved global land cover map in 2015 with 30 m resolution is available at https://doi.org/10.6084/m9.figshare.22358143.v2 (Li et al., 2022). The GLC-2015 product is organized by a total of 1507 4° × 4° geographical grids in GeoTIFF format across the world's terrestrial area. Each image of the GLC-2015 product is named as "GLC-2015_lon_lat" (lon and lat represent the longitude and latitude and of the grid's lower left corner, respectively).

6. Conclusions

GLC information at fine spatial resolution is vital for the global environment and climate studies which can capture the footprint of human activity. Resulting from the differences in classification scheme, satellite sensor data, classification algorithms and sampling strategies, the existing GLC products have

high inconsistency in some parts of the world, especially in fragmented areas and transition zones. More accurate and reliable data with accuracy improved in areas of high mapping inconsistency is very desirable. In this study, with the help of the GEE platform, we developed the GLC-2015 map by integrating multiple existing GLC maps based on the DSET. The GLC-2015 map can significantly increase the mapping accuracy and possess good recognition performance in various LC classes.

The GLC-2015 map was validated by both the global point-based samples and the global patch-based samples. Accuracy assessments show that the GLC-2015 map achieved an OA of 79.5%, a kappa coefficient of 0.757 using a total of 34,117 global point-based samples, and an OA of 83.6%, a kappa coefficient of 0.566 using a total of 201 global patch-based samples. Data inter-comparison indicated that the GLC-2015 map surpassed other three products both visually and quantitatively, by OA improvement of 14.0%-17.8% validated with the global point-based samples and 5.9%-24.5% with the global patch-based samples. Compared to other products, there are fewer misclassifications in the GLC-2015 map for most LC classes, such as forest, cropland, shrubland, and water bodies. Meanwhile, the GLC-2015 map outperformed others in terms of OA and kappa coefficient across different ecoregions and different continents. Notably, the GLC-2015 map showed better performance than others by an increment of 0.1%-0.6%. in overall accuracy for areas of low inconsistency, 19.3%-28.0% for areas of moderate inconsistency, and 27.5%-29.7% for areas of high inconsistency. In addition, the mapping results obtained by the DSET surpassed other data fusion methods with OA improvement of 7.4%-7.7% via the global point-based samples and 3.5%-11.8% via the global patch-based samples. Therefore, it can be concluded that the GLC-2015 map is a robust and reliable map that can significantly improve mapping accuracy compared to previous GLC products and mapping results from other common data fusion methods.

Author contributions

XL and XX conceived the research. BL and XX designed and carried out the experiments. QS and DH provided data. BL wrote the original manuscript. XX, HZ and YC reviewed the writing.

Competing interests

The authors declare that they have no conflict of interest.

Financial support

This research has been supported by the National Key Research & Development Program of China (Grant No. 2019YFA0607203), the National Natural Science Foundation of China (Grant No. 42001326, 42171409), and the Natural Science Foundation of Guangdong Province of China (Grant No. 2022A1515012207).

References

Ban, Y., Gong, P., and Giri, C.: Global land cover mapping using Earth observation satellite data: Recent progresses and challenges, ISPRS J. Photogramm., 103, 1-6, https://doi.org/10.1016/j.isprsjprs.2015.01.001, 2015.

Bartholomé, E. and Belward, A. S.: GLC2000: A new approach to global land cover mapping from Earth observation data, Int. J. Remote Sens., 26, 1959-1977, https://doi.org/10.1080/01431160412331291297, 2005.

Bounoua, L., DeFries, R., Collatz, G. J., Sellers, P., and Khan, H.: Effects of land cover conversion on surface climate, Clim. Change, 52, 29-64, https://doi.org/10.1023/A:1013051420309, 2002.

Bunting, P., Rosenqvist, A., Lucas, R. M., Rebelo, L. M., Hilarides, L., Thomas, N., Hardy, A., Itoh, T., Shimada, M., and Finlayson, C. M.: The Global Mangrove Watch—A new 2010 global baseline of mangrove extent, Remote Sens., 10, https://doi.org/10.3390/rs10101669, 2018.

Chapin, F. S. I., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., Hooper, D. U., Lavorel, S., Sala, O. E., Hobbie, S. E., Mack, M. C., and Díaz, S.: Consequences of changing biodiversity, Nature, 405, 234-242, https://doi.org/10.1038/35012241, 2000.

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

Chen, T. M. and Venkataramanan, V.: Dempster-Shafer theory for intrusion detection in ad hoc networks, IEEE Internet comput., 9, 35-41, https://doi.org/10.1109/MIC.2005.123, 2005.

Clinton, N., Yu, L., and Gong, P.: Geographic stacking: Decision fusion to increase global land cover map accuracy, ISPRS J. Photogramm., 103, 57-65, https://doi.org/10.1016/j.isprsjprs.2015.02.010, 2015.

Land Cover CCI: Product User Guide Version 2: https://www.esa-landcover-cci.org/?q=webfm_send/84, last access: 21 January 2022.

DeFries, R. S., Houghton, R. A., Hansen, M. C., Field, C. B., Skole, D., and Townshend, J.: Carbon emissions from tropical deforestation and regrowth based on satellite observations for the 1980s and 1990s, Proc. Natl. Acad. Sci. U.S.A., 99, 14256, https://doi.org/10.1073/pnas.182560099, 2002.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, I. C., Ramankutty, N., and Snyder, P. K.: Global Consequences of Land Use, Science, 309, 570-574, https://doi.org/10.1126/science.1111772, 2005.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., and Huang, X.: MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets, Remote Sens. Environ., 114, 168-182, https://doi.org/10.1016/j.rse.2009.08.016, 2010.

Fritz, S., You, L., Bun, A., See, L., McCallum, I., Schill, C., Perger, C., Liu, J., Hansen, M., and Obersteiner, M.: Cropland for sub-Saharan Africa: A synergistic approach using five land cover data sets, Geophy. Res. Lett., 38, L04404, https://doi.org/10.1029/2010GL046213, 2011.

Gao, Y., Liu, L., Zhang, X., Chen, X., Mi, J., and Xie, S.: Consistency analysis and accuracy assessment of three global 30 m land-cover products over the European Union using the LUCAS dataset, Remote Sens., 12, 3479, https://doi.org/10.3390/rs12213479, 2020.

Gengler, S. and Bogaert, P.: Combining land cover products using a minimum divergence and a Bayesian data fusion approach, Int. J. Geogr. Inf. Sci., 32, 806-826, https://doi.org/10.1080/13658816.2017.1413577, 2018.

Giri, C., Zhu, Z., and Reed, B.: A comparative analysis of the Global Land Cover 2000 and MODIS land cover datasets, Remote Sens. Environ., 94, 123-132, https://doi.org/10.1016/j.rse.2004.09.005, 2005.

Giri, C., Pengra, B., Long, J., and Loveland, T. R.: Next generation of global land cover characterization, mapping, and monitoring, Int. J. Appl. Earth Observ., 25, 30-37, https://doi.org/10.1016/j.jag.2013.03.005, 2013.

Gómez, C., White, J. C., and Wulder, M. A.: Optical remotely sensed time series data for land cover classification: A review, ISPRS J. Photogramm., 116, 55-72, https://doi.org/10.1016/j.isprsjprs.2016.03.008, 2016.

Gong, P.: Remote sensing of environmental change over China: A review, Sci. Bull., 57, 2793-2801, https://doi.org/10.1007/s11434-012-5268-y, 2012.

Gong, P., Yu, L., Li, C., Wang, J., Liang, L., Li, X., Ji, L., Bai, Y., Cheng, Y., and Zhu, Z.: A new research paradigm for global land cover mapping, Ann. GIS, 22, 87-102, https://doi.org/10.1080/19475683.2016.1164247, 2016.

Gong, P., Li, X., Wang, J., Bai, Y., Chen, B., Hu, T., Liu, X., Xu, B., Yang, J., Zhang, W., and Zhou, Y.: Annual maps of global artificial impervious area (GAIA) between 1985 and 2018, Remote Sens. Environ., 236, 111510, https://doi.org/10.1016/j.rse.2019.111510, 2020.

Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, H., Zhu, P., Zhao, Z., Zhang, H., Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, J., and Chen, J.: Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data, Int. J. Remote Sens., 34, 2607-2654, https://doi.org/10.1080/01431161.2012.748992, 2013.

Grekousis, G., Mountrakis, G., and Kavouras, M.: An overview of 21 global and 43 regional land-cover mapping products, Int. J. Remote Sens., 36, 5309-5335, https://doi.org/10.1080/01431161.2015.1093195, 2015.

Grimm, N. B., Faeth, S. H., Golubiewski, N. E., Redman, C. L., Wu, J., Bai, X., and Briggs, J. M.: Global change and the ecology of cities, Science, 319, 756-760, https://doi.org/10.1126/science.1150195, 2008.

Hansen, M. C., Defries, R. S., Townshend, J. R. G., and Sohlberg, R.: Global land cover classification at 1 km spatial resolution using a classification tree approach, Int. J. Remote Sens., 21, 1331-

1364, https://doi.org/10.1080/014311600210209, 2000.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., and Townshend, J. R. G.: High-resolution global maps of 21st-century forest cover change, Science, 342, 850-853, https://doi.org/10.1126/science.1244693, 2013.

Herold, M., Mayaux, P., Woodcock, C. E., Baccini, A., and Schmullius, C.: Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets, Remote Sens. Environ., 112, 2538-2556, https://doi.org/10.1016/j.rse.2007.11.013, 2008.

Hu, L., Chen, Y., Xu, Y., Zhao, Y., Yu, L., Wang, J., and Gong, P.: A 30 meter land cover mapping of China with an efficient clustering algorithm CBEST, Science China Earth Sciences, 57, 2293-2304, https://doi.org/10.1007/s11430-014-4917-1, 2014.

Huang, X., Li, J., Yang, J., Zhang, Z., Li, D., and Liu, X.: 30 m global impervious surface area dynamics and urban expansion pattern observed by Landsat satellites: From 1972 to 2019, Science China Earth Sciences, 64, 1922-1933, https://10.1007/s11430-020-9797-9, 2021.

Iwao, K., Nasahara, K. N., Kinoshita, T., Yamagata, Y., Patton, D., and Tsuchida, S.: Creation of new global land cover map with map integration, J. Geogr. Inf. Syst., 3, 160-165, https://doi.org/10.4236/jgis.2011.32013, 2011.

Jin, Q., Xu, E., and Zhang, X.: A fusion method for multisource land cover products based on superpixels and statistical extraction for enhancing resolution and improving accuracy, Remote Sens., 14, 1676, https://doi.org/10.3390/rs14071676, 2022.

Jung, M., Henkel, K., Herold, M., and Churkina, G.: Exploiting synergies of global land cover products for carbon cycle modeling, Remote Sens. Environ., 101, 534-553, https://doi.org/10.1016/j.rse.2006.01.020, 2006.

Kang, J., Wang, Z., Sui, L., Yang, X., Ma, Y., and Wang, J.: Consistency analysis of remote sensing land cover products in the tropical rainforest climate region: A case study of Indonesia, Remote Sens., 12, 1410, https://doi.org/10.3390/rs12091410, 2020.

Kim, D., Lim, C.-H., Song, C., Lee, W.-K., Piao, D., Heo, S., and Jeon, S.: Estimation of future carbon budget with climate change and reforestation scenario in North Korea, Adv. Space Res., 58, 1002-1016, https://doi.org/10.1016/j.asr.2016.05.049, 2016.

Li, B., Xu, X., Liu, X., Shi, Q., Zhuang, H., Cai, Y., and He, D.: An improved global land cover mapping in 2015 with 30 m resolution (GLC-2015) based on a multi-source product fusion approach. [dataset], https://doi.org/10.6084/m9.figshare.22358143.v2, 2022.

Li, C., Gong, P., Wang, J., Zhu, Z., Biging, G. S., Yuan, C., Hu, T., Zhang, H., Wang, Q., Li, X., Liu, X., Xu, Y., Guo, J., Liu, C., Hackman, K. O., Zhang, M., Cheng, Y., Yu, L., Yang, J., Huang, H., and Clinton, N.: The first all-season sample set for mapping global land cover with Landsat-8 data, Sci. Bull., 62, 508-515, https://doi.org/10.1016/j.scib.2017.03.011, 2017.

Liu, H., Gong, P., Wang, J., Clinton, N., Bai, Y., and Liang, S.: Annual dynamics of global land cover and its long-term changes from 1982 to 2015, Earth Syst. Sci. Data, 12, 1217-1243, https://doi.org/10.5194/essd-12-1217-2020, 2020a.

Liu, H., Gong, P., Wang, J., Wang, X., Ning, G., and Xu, B.: Production of global daily seamless data cubes and quantification of global land cover change from 1985 to 2020 - iMap World 1.0, Remote Sens. Environ., 258, 112364, https://doi.org/10.1016/j.rse.2021.112364, 2021a.

Liu, J., Kuang, W., Zhang, Z., Xu, X., Qin, Y., Ning, J., Zhou, W., Zhang, S., Li, R., Yan, C., Wu, S., Shi, X., Jiang, N., Yu, D., Pan, X., and Chi, W.: Spatiotemporal characteristics, patterns and causes of

land use changes in China since the late 1980s, Dili Xuebao/Acta Geogr. Sin., 69, 3-14, https://doi.org/10.11821/dlxb201401001, 2014.

Liu, K. and Xu, E.: Fusion and correction of multi-source land cover products based on spatial detection and uncertainty reasoning methods in Central Asia, Remote Sens., 13, 244, https://doi.org/10.3390/rs13020244, 2021.

Liu, L., Zhang, X., Gao, Y., Chen, X., Shuai, X., and Mi, J.: Finer-resolution mapping of global land cover: Recent developments, consistency analysis, and prospects, Journal of Remote Sensing, 2021, 5289697, https://doi.org/10.34133/2021/5289697, 2021b.

Liu, X., Huang, Y., Xu, X., Li, X., Li, X., Ciais, P., Lin, P., Gong, K., Ziegler, A. D., Chen, A., Gong, P., Chen, J., Hu, G., Chen, Y., Wang, S., Wu, Q., Huang, K., Estes, L., and Zeng, Z.: High-spatiotemporal-resolution mapping of global urban change from 1985 to 2015, Nature Sustainability, 3, 564-570, https://doi.org/10.1038/s41893-020-0521-x, 2020b.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., and Merchant, J. W.: Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data, Int. J. Remote Sens., 21, 1303-1330, https://doi.org/10.1080/014311600210191, 2000.

Ludwig, C., Walli, A., Schleicher, C., Weichselbaum, J., and Riffler, M.: A highly automated algorithm for wetland detection using multi-temporal optical satellite data, Remote Sens. Environ., 224, 333-351, https://doi.org/10.1016/j.rse.2019.01.017, 2019.

Mayaux, P., Bartholomé, E., Fritz, S., and Belward, A.: A new land-cover map of Africa for the year 2000, J. Biogeogr., 31, 861-877, https://doi.org/10.1111/j.1365-2699.2004.01073.x, 2004.

McCallum, I., Obersteiner, M., Nilsson, S., and Shvidenko, A.: A spatial comparison of four satellite derived 1km global land cover datasets, Int. J. Appl. Earth Observ., 8, 246-255, https://doi.org/10.1016/j.jag.2005.12.002, 2006.

Meng, Z., Dong, J., Ellis, E. C., Metternicht, G., Qin, Y., Song, X.-P., Löfqvist, S., Garrett, R. D., Jia, X., and Xiao, X.: Post-2020 biodiversity framework challenged by cropland expansion in protected areas, Nature Sustainability, https://doi.org/10.1038/s41893-023-01093-w, 2023.

Meyer, M. F., Labou, S. G., Cramer, A. N., Brousil, M. R., and Luff, B. T.: The global lake area, climate, and population dataset, Sci. Data, 7, 174, https://doi.org/10.1038/s41597-020-0517-4, 2020.

Moody, A. and Woodcock, C.: Scale-dependent errors in the estimation of land-cover proportions: Implications for global land-cover datasets, Photogramm. Eng. Remote Sens., 60, 585-594, 1994.

Pekel, J. F., Cottam, A., Gorelick, N., and Belward, A. S.: High-resolution mapping of global surface water and its long-term changes, Nature, 540, 418-422, https://doi.org/10.1038/nature20584, 2016.

Pengra, B. W., Stehman, S. V., Horton, J. A., Dockter, D. J., Schroeder, T. A., Yang, Z., Cohen, W. B., Healey, S. P., and Loveland, T. R.: Quality control and assessment of interpreter consistency of annual land cover reference data in an operational national monitoring program, Remote Sens. Environ., 238, 111261, https://doi.org/10.1016/j.rse.2019.111261, 2020.

Razi, S., Karami Mollaei, M. R., and Ghasemi, J.: A novel method for classification of BCI multi-class motor imagery task based on Dempster–Shafer theory, Inf. Sci., 484, 14-26, https://doi.org/10.1016/j.ins.2019.01.053, 2019.

Rottensteiner, F., Trinder, J. C., Clode, S., and Kubik, K.: Using the Dempster-Shafer method for the fusion of LIDAR data and multi-spectral images for building detection, Inform. Fusion., 6, 283-300, https://doi.org/10.1016/j.inffus.2004.06.004, 2005.

Running, S. W.: Ecosystem disturbance, carbon, and climate, Science, 321, 652-653, https://doi.org/10.1126/science.1159607, 2008.

Schewe, J., Gosling, S. N., Reyer, C., Zhao, F., Ciais, P., Elliott, J., Francois, L., Huber, V., Lotze, H. K., Seneviratne, S. I., van Vliet, M. T. H., Vautard, R., Wada, Y., Breuer, L., Büchner, M., Carozza, D. A., Chang, J., Coll, M., Deryng, D., de Wit, A., Eddy, T. D., Folberth, C., Frieler, K., Friend, A. D., Gerten, D., Gudmundsson, L., Hanasaki, N., Ito, A., Khabarov, N., Kim, H., Lawrence, P., Morfopoulos, C., Müller, C., Müller Schmied, H., Orth, R., Ostberg, S., Pokhrel, Y., Pugh, T. A. M., Sakurai, G., Satoh, Y., Schmid, E., Stacke, T., Steenbeek, J., Steinkamp, J., Tang, Q., Tian, H., Tittensor, D. P., Volkholz, J., Wang, X., and Warszawski, L.: State-of-the-art global models underestimate impacts from climate extremes, Nat. Commun., 10, 1005, https://doi.org/10.1038/s41467-019-08745-6, 2019.

See, L., Schepaschenko, D., Lesiv, M., McCallum, I., Fritz, S., Comber, A., Perger, C., Schill, C., Zhao, Y., Maus, V., Siraj, M. A., Albrecht, F., Cipriani, A., Vakolyuk, M. y., Garcia, A., Rabia, A. H., Singha, K., Marcarini, A. A., Kattenborn, T., Hazarika, R., Schepaschenko, M., van der Velde, M., Kraxner, F., and Obersteiner, M.: Building a hybrid land cover map with crowdsourcing and geographically weighted regression, ISPRS J. Photogramm., 103, 48-56, https://doi.org/10.1016/j.isprsjprs.2014.06.016, 2015.

Shafizadeh-Moghadam, H., Minaei, M., Feng, Y., and Pontius, R. G.: GlobeLand30 maps show four times larger gross than net land change from 2000 to 2010 in Asia, Int. J. Appl. Earth Observ., 78, 240-248, https://doi.org/10.1016/j.jag.2019.01.003, 2019.

Song, X., Hansen, M. C., Stehman, S. V., Potapov, P. V., Tyukavina, A., Vermote, E. F., and Townshend, J. R.: Global land change from 1982 to 2016, Nature, 560, 639-643, https://doi.org/10.1038/s41586-018-0411-9, 2018.

Sun, B., Chen, X., and Zhou, Q.: Uncertainty assessment of GlobeLand30 land cover data set over central Asia, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., 41, 1313, https://doi.org/10.5194/isprs-archives-XLI-B8-1313-2016, 2016.

Teluguntla, P., Thenkabail, P. S., Oliphant, A., Xiong, J., Gumma, M. K., Congalton, R. G., Yadav, K., and Huete, A.: A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform, ISPRS J. Photogramm., 144, 325-340, https://doi.org/10.1016/j.isprsjprs.2018.07.017, 2018.

Verburg, P. H., Neumann, K., and Nol, L.: Challenges in using land use and land cover data for global change studies, Glob. Change Biol., 17, 974-989, https://doi.org/10.1111/j.1365-2486.2010.02307.x, 2011.

Verburg, P. H., Mertz, O., Erb, K.-H., Haberl, H., and Wu, W.: Land system change and food security: towards multi-scale land system solutions, Curr. Opin. Environ. Sustain., 5, 494-502, https://doi.org/10.1016/j.cosust.2013.07.003, 2013.

Wickham, J., Stehman, S. V., Sorenson, D. G., Gass, L., and Dewitz, J. A.: Thematic accuracy assessment of the NLCD 2016 land cover for the conterminous United States, Remote Sens. Environ., 257, 112357, https://doi.org/10.1016/j.rse.2021.112357, 2021.

Wu, J., Wang, X., Zhong, B., Yang, A., Jue, K., Wu, J., Zhang, L., Xu, W., Wu, S., Zhang, N., and Liu, Q.: Ecological environment assessment for Greater Mekong Subregion based on Pressure-State-Response framework by remote sensing, Ecol. Indic., 117, 106521, https://doi.org/10.1016/j.ecolind.2020.106521, 2020.

Wulder, M. A., Li, Z., Campbell, E. M., White, J. C., Hobart, G., Hermosilla, T., and Coops, N. C.: A

1098   national assessment of wetland status and trends for Canada's forested ecosystems using 33
1099   years of earth observation satellite data, Remote Sens., 10, 1623,
1100   https://doi.org/10.3390/rs10101623, 2018.
1101 Xiong, J., Thenkabail, P. S., Tilton, J. C., Gumma, M. K., Teluguntla, P., Oliphant, A., Congalton, R. G.,
1102   Yadav, K., and Gorelick, N.: Nominal 30-m cropland extent map of continental Africa by
1103   integrating pixel-based and object-based algorithms Using Sentinel-2 and Landsat-8 data on
1104   Google Earth Engine, Remote Sens., 9, 1065, https://doi.org/10.3390/rs9101065, 2017.
1105 Xu, G., Zhang, H., Chen, B., Zhang, H., Yan, J., Chen, J., Che, M., Lin, X., and Dou, X.: A Bayesian
1106   based method to generate a synergetic land-cover map from existing land-cover products.,
1107   Remote Sens., 6, 5589-5613, https://doi.org/10.3390/rs606558910.3390/rs6065589, 2014.
1108 Xue, J., Wang, Y., Teng, H., Wang, N., Li, D., Peng, J., Biswas, A., and Shi, Z.: Dynamics of vegetation
1109   greenness and lits response to climate change in Xinjiang over the past two decades, Remote
1110   Sens., 13, 4063, https://doi.org/10.3390/rs13204063, 2021.
1111 Yang, J. and Huang, X.: The 30 m annual land cover dataset and its dynamics in China from 1990
1112   to 2019, Earth Syst. Sci. Data, 13, 3907-3925, https://doi.org/10.5194/essd-13-3907-2021, 2021.
1113 Yang, J., Gong, P., Fu, R., Zhang, M., Chen, J., Liang, S., Xu, B., Shi, J., and Dickinson, R.: The role of
1114   satellite remote sensing in climate change studies, Nat. Clim. Chang., 3, 875-883,
1115   https://doi.org/10.1038/nclimate1908, 2013.
1116 Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S. M., Case, A., Costello, C., Dewitz, J.,
1117   Fry, J., Funk, M., Granneman, B., Liknes, G. C., Rigge, M., and Xian, G.: A new generation of the
1118   United States National Land Cover Database: Requirements, research priorities, design, and
1119   implementation strategies, ISPRS J. Photogramm., 146, 108-123,
1120   https://doi.org/10.1016/j.isprsjprs.2018.09.006, 2018.
1121 Yang, Y., Xiao, P., Feng, X., and Li, H.: Accuracy assessment of seven global land cover datasets
1122   over China, ISPRS J. Photogramm., 125, 156-173, https://doi.org/10.1016/j.isprsjprs.2017.01.016,
1123   2017.
1124 Zhang, C., Dong, J., and Ge, Q.: Quantifying the accuracies of six 30-m cropland datasets over
1125   China: A comparison and evaluation analysis, Comput. Electron. Agric., 197, 106946,
1126   https://doi.org/10.1016/j.compag.2022.106946, 2022.
1127 Zhang, M., Ma, M., De Maeyer, P., and Kurban, A.: Uncertainties in classification system conversion
1128   and an analysis of inconsistencies in global land cover products, ISPRS Int. J. Geo Inf., 6, 112,
1129   https://doi.org/10.3390/ijgi6040112, 2017.
1130 Zhang, X., Liu, L., Chen, X., Gao, Y., Xie, S., and Mi, J.: GLC_FCS30: global land-cover product with
1131   fine classification system at 30 m using time-series Landsat imagery, Earth Syst. Sci. Data, 13,
1132   2753-2776, https://doi.org/10.5194/essd-13-2753-2021, 2021.
1133 Zhang, X., Liu, L., Zhao, T., Chen, X., Lin, S., Wang, J., Mi, J., and Liu, W.: GWL_FCS30: a global 30 m
1134   wetland map with a fine classification system using multi-sourced and time-series remote
1135   sensing imagery in 2020, Earth Syst. Sci. Data, 15, 265-293, https://doi.org/10.5194/essd-15-
1136   265-2023, 2023.
1137 Zhao, J., Yu, L., Liu, H., Huang, H., Wang, J., and Gong, P.: Towards an open and synergistic
1138   framework for mapping global land cover, PeerJ, 9, e11877, https://doi.org/10.7717/peerj.11877,
1139   2021.
1140 Zheng, W., Liu, Y., Yang, X., and Fan, W.: Spatiotemporal variations of forest vegetation phenology
1141   and its response to climate change in northeast China, Remote Sens., 14, 2909,

1142   https://doi.org/10.3390/rs14122909, 2022.

1143