Thank you for the comments and suggestions. These comments were very helpful for revising and improving our paper. We have responded to the comments point by point and made the detailed revisions embedded in the manuscript with the line numbers indicated in the responses.

Comment 1: This paper developed a 30-m resolution dataset of China's urban impervious surface area and green space fraction from 2000 to 2018. The fraction information of impervious and green space is revealed from remotely sensed indicators. I have some comments on the adopted approaches and presented results.

Main comments (1) It is difficult to derive the logistic regression model directly since most of the available observations are limited in this study (i.e., only five observations). Why not use continuous observations (i.e., annual) to fit the logistic regression model.

Response: Thank you for your comments. Recently, we published a 2020 annual report by Global Ecosystems and Environment Observation Analysis Research Cooperation (http://www.chinageoss.org/geoarc/2020/) through cooperation between the Global Earth Observation System of Systems (GEOSS) and the National Remote Sensing Center of China at the Ministry of Science and Technology. We developed a set of new algorithms to retrieve the UIS and UGS fractions using sub-pixel decomposition method through random forest algorithm using Google Earth Engine (GEE) platform. We improved the methods on mapping UIS and UGS fractions and updated the datasets of CLUD-Urban product.

Changes in manuscript: We rewrote the method on the retrieval of UIS and UGS fractions below in L170-205.

Comment 2: Retrieving the ISA information from the NDVI directly seems not reasonable. First, the maximum NDVI used in this study may be fluctuated over years and across spaces. Second, for the built-up areas in arid regions, the proposed approach of estimating the impervious surface information from the NDVI is not reliable. This is an issue that needs to be well addressed.

Response: Thank you for your comments. We updated an algorithm to acquire the UISA and UGS fractions. We found the new method and datasets have a high reliability to address those issues.

Changes in manuscript: We updated the method on the retrieval of UGS fraction in Line 170-205.

5 Mapping UIS and UGS fractions using GEE platform

5.1 Collection of training samples

The training samples of UIS and UGS fractions are a pivotal input parameter in random forest model for mapping national settlement and vegetation fraction. In light of large discrepancies among UIS and UGS composites in different climate zones with various geographical and social economic conditions, we collected a total of 2,570 samples from randomly selected cities in different climate zones (Schneider et al. 2010) (Fig. 5). Here we also refer to the existing UIS dataset to acquire samples with 10% intervals of the ISA fraction, and those samples primarily distributed in the homogeneous UIS or UGS areas, which might provide more effective samples and decrease the impact of imagery mismatch. The samples of UIS and UGS covered with diversified types, including buildings, roads and squares, and grass, trees from parks, road and residential green spaces. The UIS and UGS percentages were interpreted within each sample using Google Earth images (Fig. 5b1-b4). Finally, the training samples in 2000, 2005, 2010, 2015 and 2018 were used for training the random forest model, respectively.

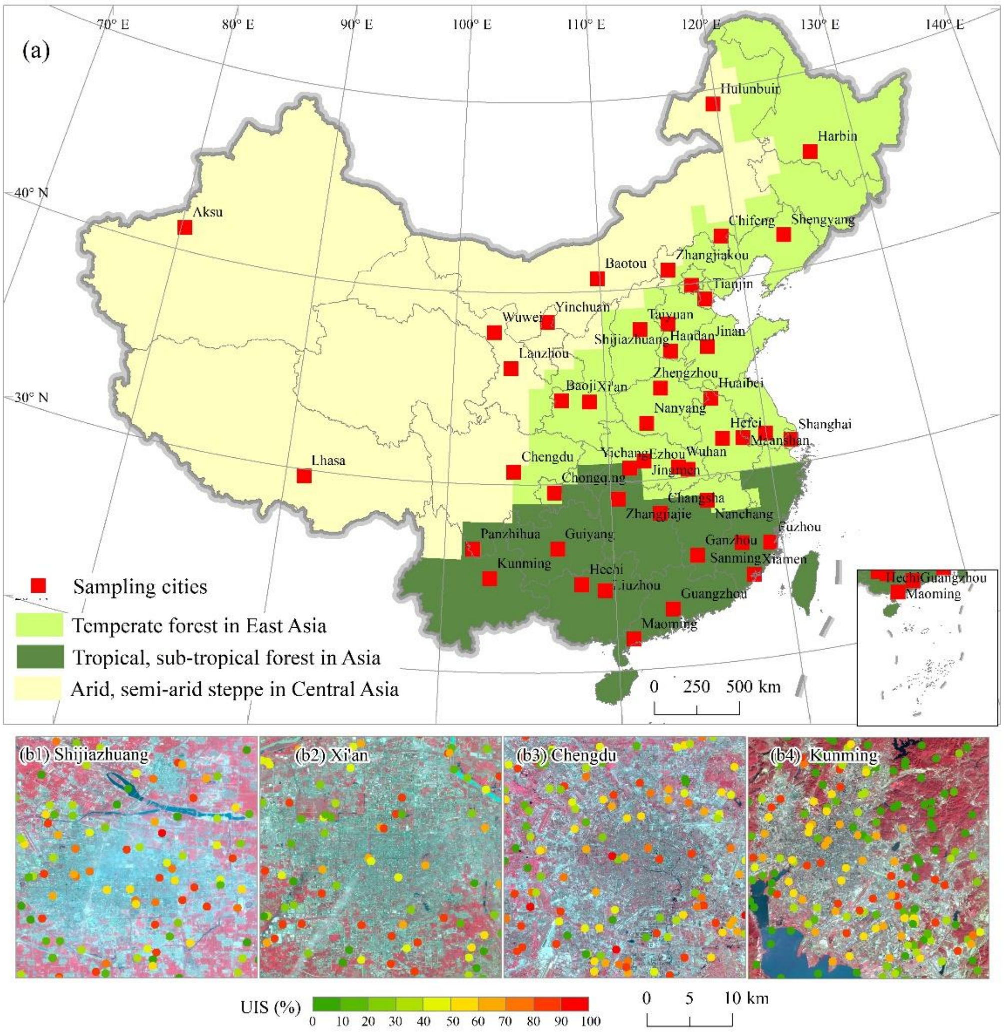

Figure 5: Distribution of sampling cities in China and training samples in selected cities. (The images were provided by Geospatial Data Cloud site, Computer Network Information Center, Chinese Academy of Sciences (http://www.gscloud.cn). The administrative boundaries were provided by National Geomatics Center of China (http://www.webmap.cn))

5.2 Retrieval of settlement and vegetation fractions using random forest model

Many previous studies have indicated that random forest is more effective and accurate in classifying urban land types than other machine learning approaches such as support vector machine (SVM) and artificial neural network (ANN) (Zhang et al., 2020). Random forest exhibits a strong capacity in processing high-dimensional datasets and has been successfully applied to mapping global ISA at 30-m resolution (Zhang et al., 2020). In this research, we proposed a strategy to acquire the settlement and vegetation percentage at pixel scale using the advantage of random forest and big-data processing based on GEE platform.

According to sixteen global urban ecoregions based on temperature, precipitation, topographic conditions and social economic factors (Schneider et al. 2010), China has three urban ecoregions. In each urban ecoregion, the annual maximum NDVI, and spectral bands in Landsat TM/ETM+/OLI, and the slope index derived from SRTM DEM with 30-m resolution were selected as the input parameters to run random forest model. The Landsat images were from January 1 to December 31 of each baseline year. The annual maximum NDVI ($NDVI_{max}$) was retrieved using equation (1):

$$NDVI_{max} = \max(NDVI_1, NDVI_2, \cdots, NDVI_i) \quad (1)$$

where $NDVI_i$ is the NDVI value of the i^{th} image. Individual NDVI was calculated from Landsat images in the period between January 1 to December 31 and all images were collected using GEE (Gorelick et al., 2017).

In GEE platform, the settlement and vegetation fractions were calculated for each urban ecoregion through using the training parametrizations. The lawn, forest or their mosaicked areas were selected as input samples in mapping UGS. A post-processing was implemented to remove the pixels with NDVI values of greater than 0.5 or DEM slope values of greater than 15°. In arid and semi-arid areas, the enhanced bare soil index (EBSI) was utilized to separate UIS from bare soils (As-syakur et al., 2012; Li et al., 2019). As a result, the settlement and vegetation fractions with 30 m x 30 m in 2000, 2005, 2010, 2015 and 2018 were generated for developing CLUD-Urban product (Fig. 6).

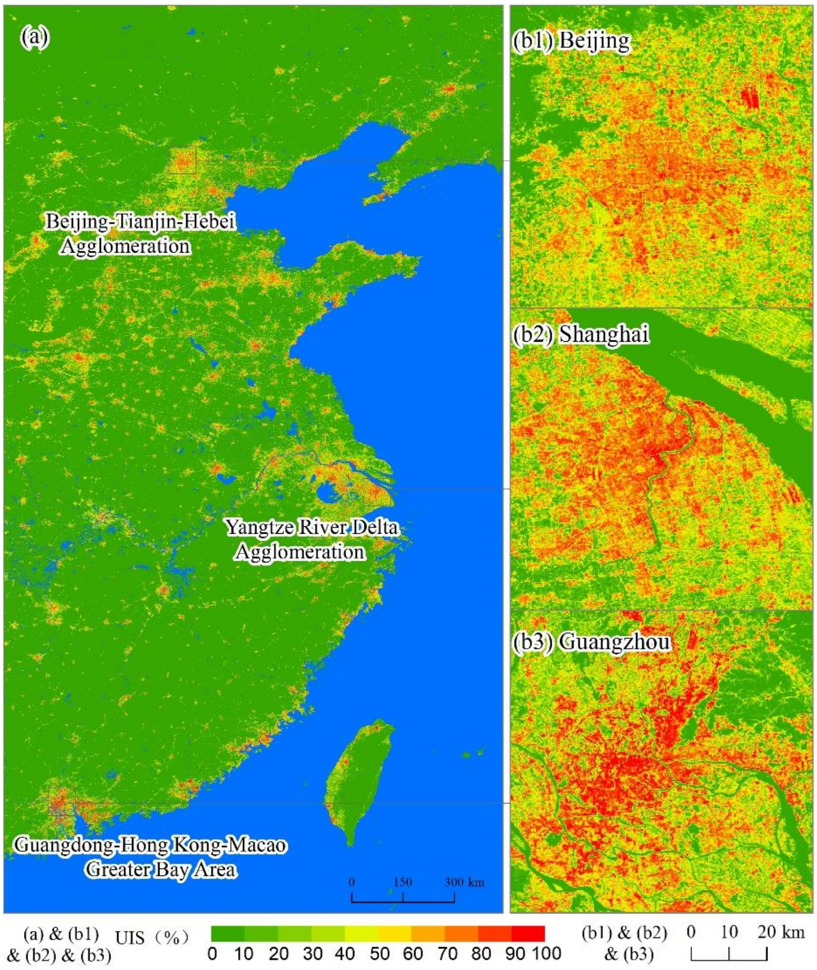

Figure 6: **Distribution of sampling cities in China and training samples in selected cities. (The administrative boundaries and residential points information were provided by National Geomatics Center of China (http://www.webmap.cn))**

5.3 Mapping of UIS and UGS fractions

The settlement and vegetation fractions with 1ºX1ºgrid of each period were downloaded from GEE platform. In ARCGIS 10.0 software, the settlement and vegetation layers were merged respectively at provincial scale with 30 m X 30 m. The national UIS and UGS fractions with 30 m X 30 m resolution in 2000, 2005, 2010, 2015 and 2018 were produced through overlaying the urban boundaries of CLUD with settlement and vegetation fractions, respectively (Fig. 7, Fig. 8 and Fig. 9).

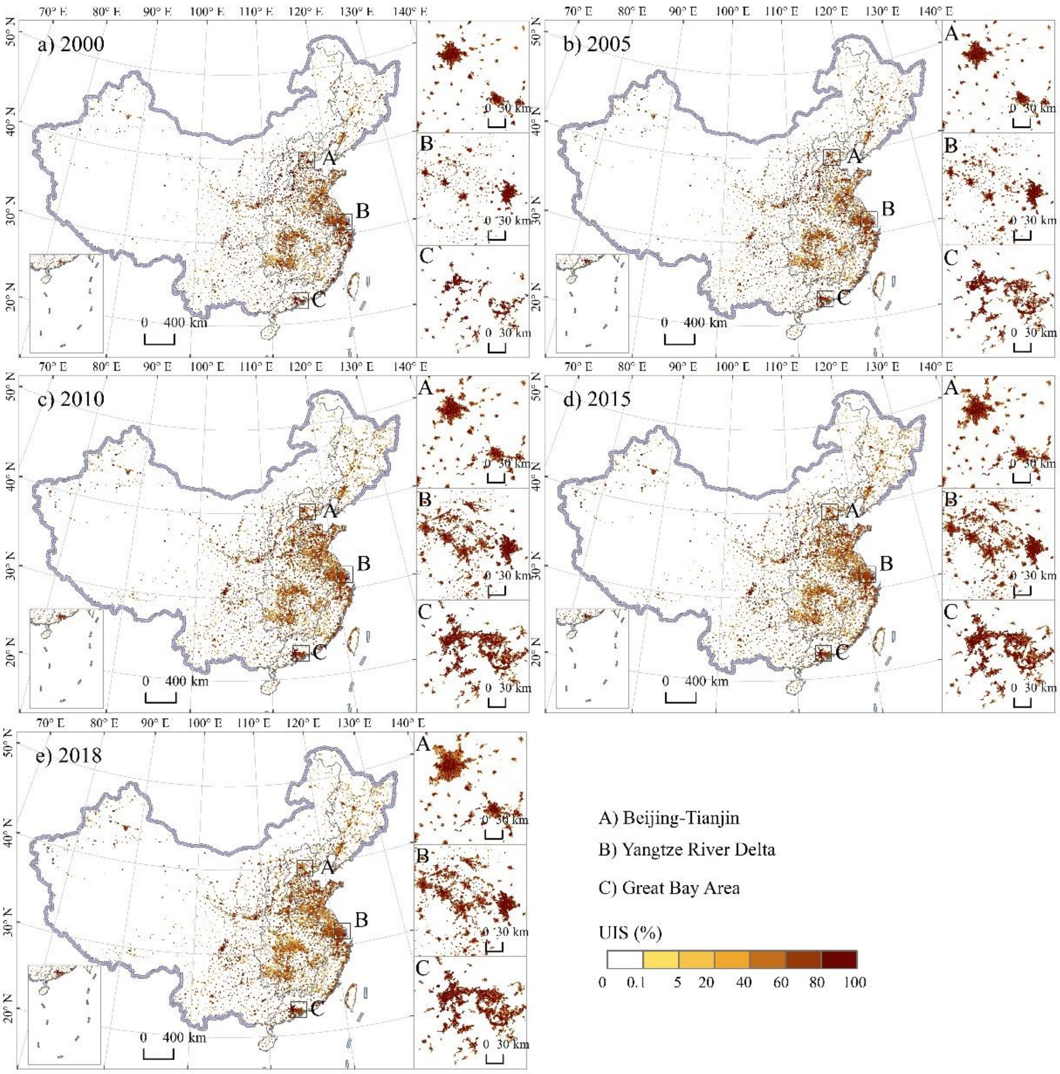

Figure 7: **Spatial distribution of urban impervious surface (UIS) in 2000–2018 across China. (The administrative boundaries were provided by National Geomatics Center of China (http://www.webmap.cn))**

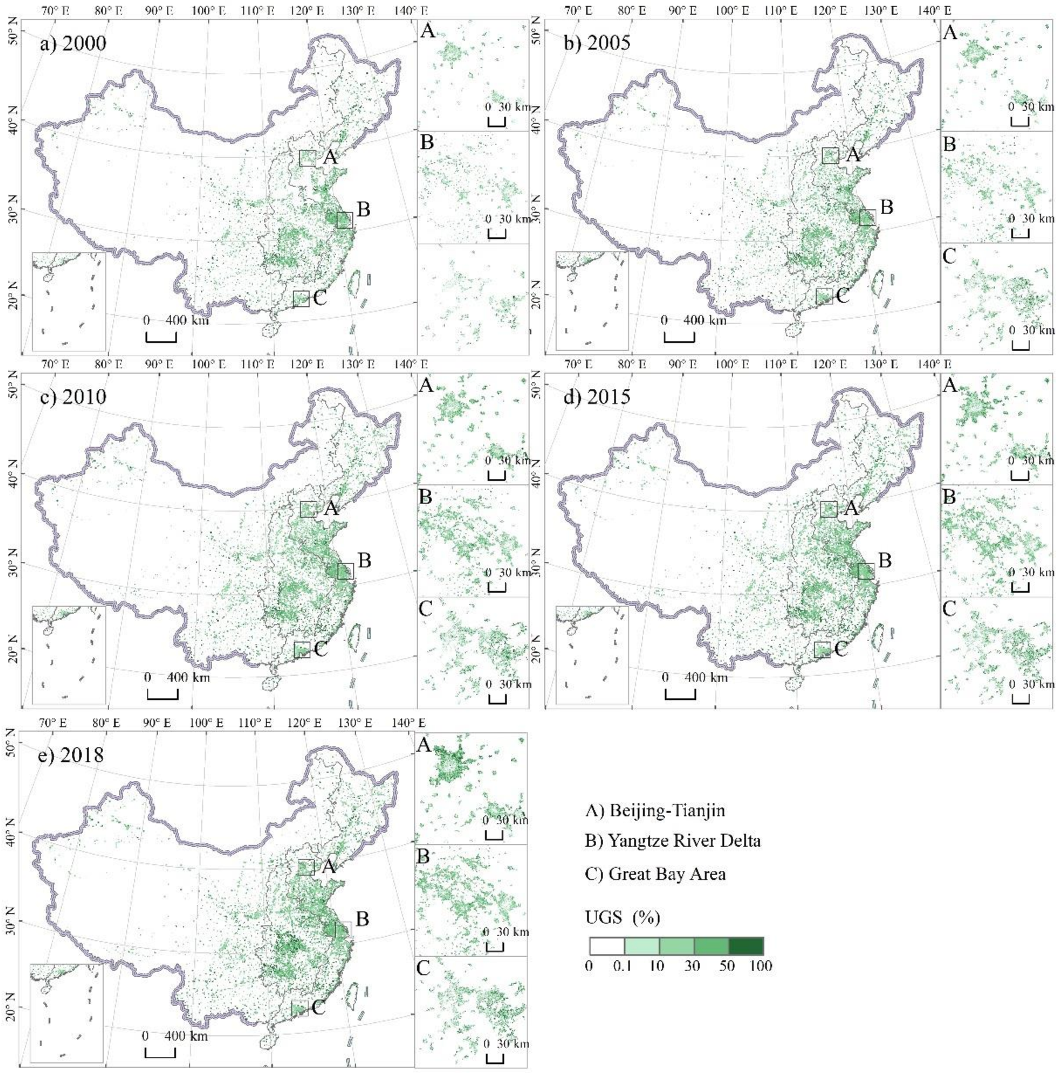

Figure 8: **Spatial distribution of urban green space (UGS) in 2000–2018 across China. (The administrative boundaries were provided by National Geomatics Center of China (http://www.webmap.cn))**

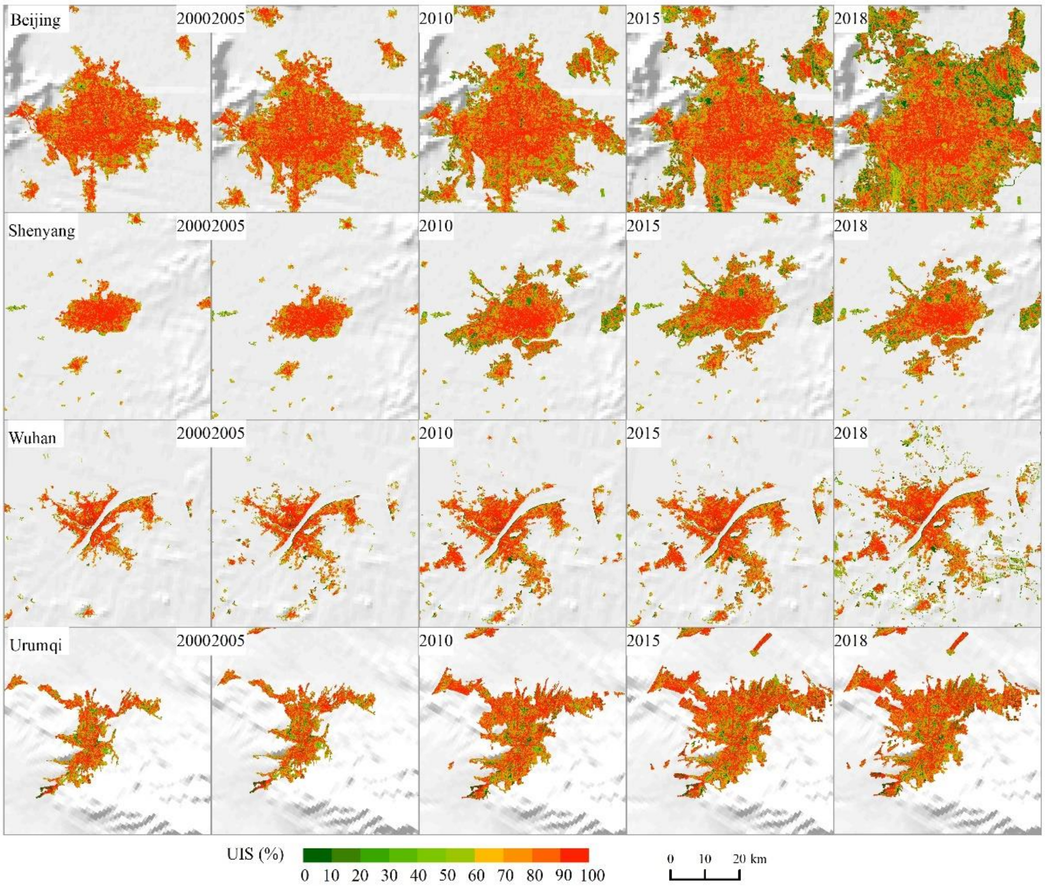

Figure 9: The change of urban impervious surface (UIS) in selected cities from coastal, central, eastern and western zones from 2000 to 2018. (DEM dataset was downloaded from SRTM 90 m Digital Elevation Data (http://srtm.csi.cgiar.org/))

Comment 3: The comparison of urban areas with other products is necessary, including the accuracy and urban area. It is inadequate if only presenting these comparable figures here directly.

Response: Thank you for your comments. Currently, the mainstream datasets on mapping global or China's urban land use/cover focus on the urban boundaries or impervious surface areas. We conducted a comparison on the different datasets and provided on the associated analysis.

Changes in manuscript: Here we added a section "7.3 Comparisons of the CLUD-Urban product with other datasets" to compare the accuracy between CLUD-Urban and the existing dataset in L270-285.

Comment 4: A similar approach to estimate the fraction of green space is not reasonable also. I am wondering if this linear relationship can well estimate the green space.

Response: Thank you for your comments. In the new dataset, we acquired the UGS fraction through random forest algorithm using Google Earth Engine (GEE) platform. We added the section to elucidate the producing flow on the new version CLUD-Urban dataset.

Changes in manuscript: We rewrote the method on the retrieval of UIS and UGS fractions below in L170-205.

Comment 5: Minor comments: Page 3, Line 78-80: how to harmonize the spectral bands between Landsat and HJ (or CBERS-1)? Page 3, Line 95: the author mentioned that most classifications exclude the green space, which is not accurate. Green space is a kind of definition from land use, which is consists of trees, shrubs, grasses etc., which are general land cover types. Page 4, Line 100: the definition of new and old urban lands from their colors still needs more evidence. Visually, the new built-up areas such as residential which has similar layout and materials may be similar to the old urban lands Page 8, Line 230: there is a repeated reference (Dong et al., 2017).

Response: Thank you for your comments.

Changes in manuscript: We revised the manuscript on the data sources and pre-processing in L120-125. We rewrote the sentence in Line 95. The definition of new and old urban lands in Landsat image isn't essential, so we replaced this image by a new image in Fig. 2. We removed the repeated reference in L370.