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: The authors present multi-year maps of urban imperviousness and greenness of China, which were estimated based on hand-drawn urban boundary and the relationship between vegetation greenness and surface imperviousness. Despite the data might be valuable to a variety of urban-related applications, there are many uncertainties remain. As a data set, these uncertainties should be clearly addressed so that users could better use it. First, using NDVI as the only indicator to estimate surface imperviousness is problematic. The NDVI-based method would overestimate the extent of impervious surfaces because of their similar characteristics as some land uses/covers on NDVI images, especially bare ground. This is especially true in most Chinese cities as they have seen substantial expansions during the study period and the extent of bare ground cannot be ignored. Second, calibration of NDVI-ISA relationship is not clear in many aspects. For example, how was ISA reference measured for model calibration? What was the performance of region averaged model compared to city-specific ones? Was the model calibrated once and applied through time or annually?

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. In newly developed CLUD-Urban product, we adopted the advantage of high accuracy and long-time series in mapping urban land from CLUD. We also utilized the highly efficient computation and large storage capacities on GEE platform. In mapping CLUD-Urban product, we proposed to quantitively retrieve the UIS and UGS fractions using random forest. The new CLUD-Urban product exhibits a high accuracy and reliability in delineating urban land surface property. Therefore, we uploaded the new version datasets on national UIS and UGS fractions dataset with 30m resolution in 2000, 2005, 2010, 2015 and 2018.

Changes in manuscript: We rewrote the fifth part on "5. Method of mapping UIS and UGS fractions using GEE platform", including three sub-sections: the collection of training samples, retrieval of settlement and vegetation fractions using random forest, and mapping of UIS and UGS fractions in L170-205.

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.

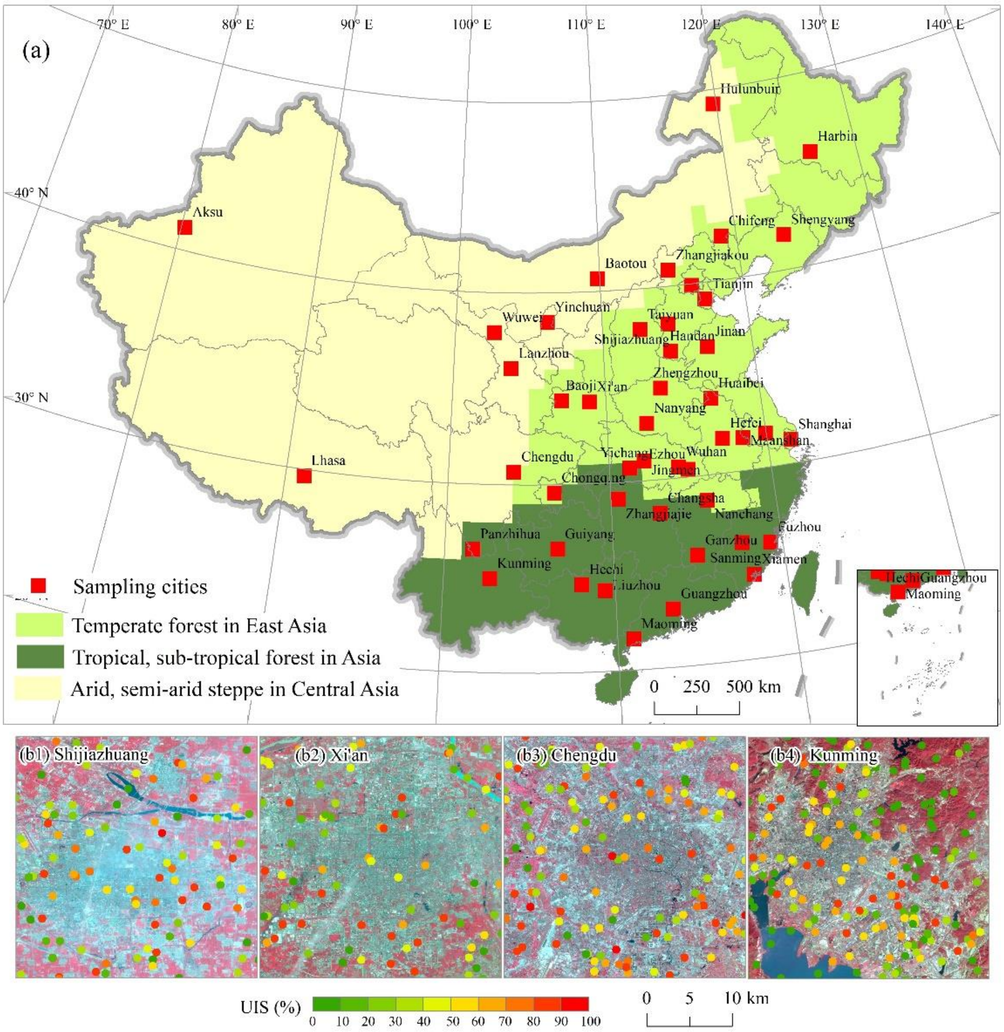

Figure 5: Distribution of sampling cities in China and training samples in selected cities. (The images were provided by Geospatial Data Cloud site, Computer Network Information Center, Chinese Academy of Sciences (http://www.gscloud.cn). The administrative boundaries were provided by National Geomatics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

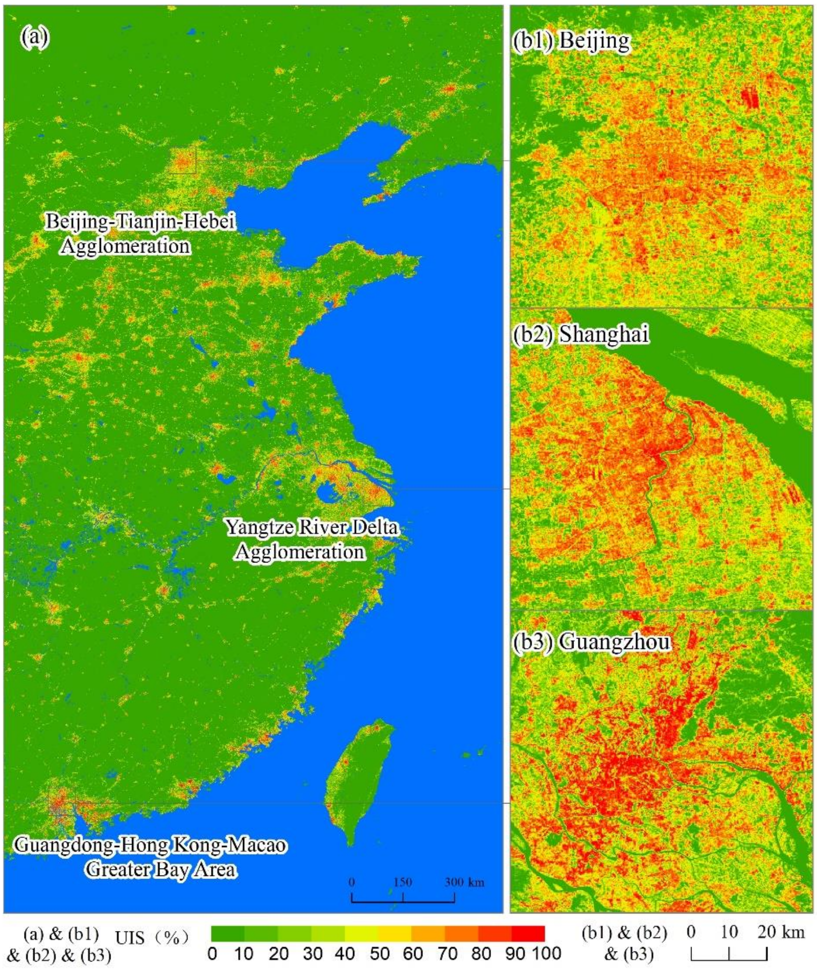

Figure 6: Distribution of sampling cities in China and training samples in selected cities. (The

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 1°X1°grid 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).

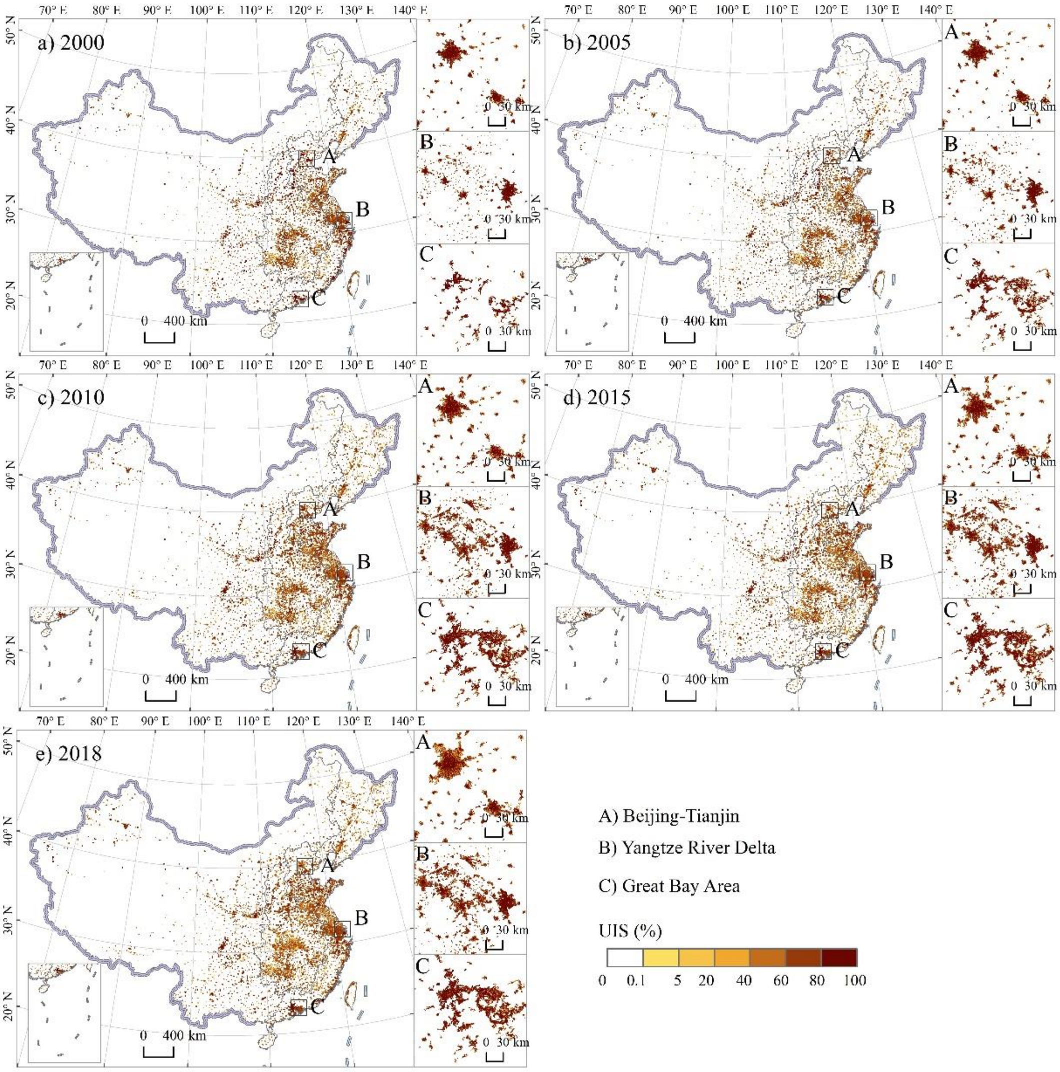

Figure 7: **Spatial distribution of urban impervious surface (UIS) in 2000–2018 across China. (The administrative boundaries were provided by National Geomatics Center of China (http://www.webmap.cn))**

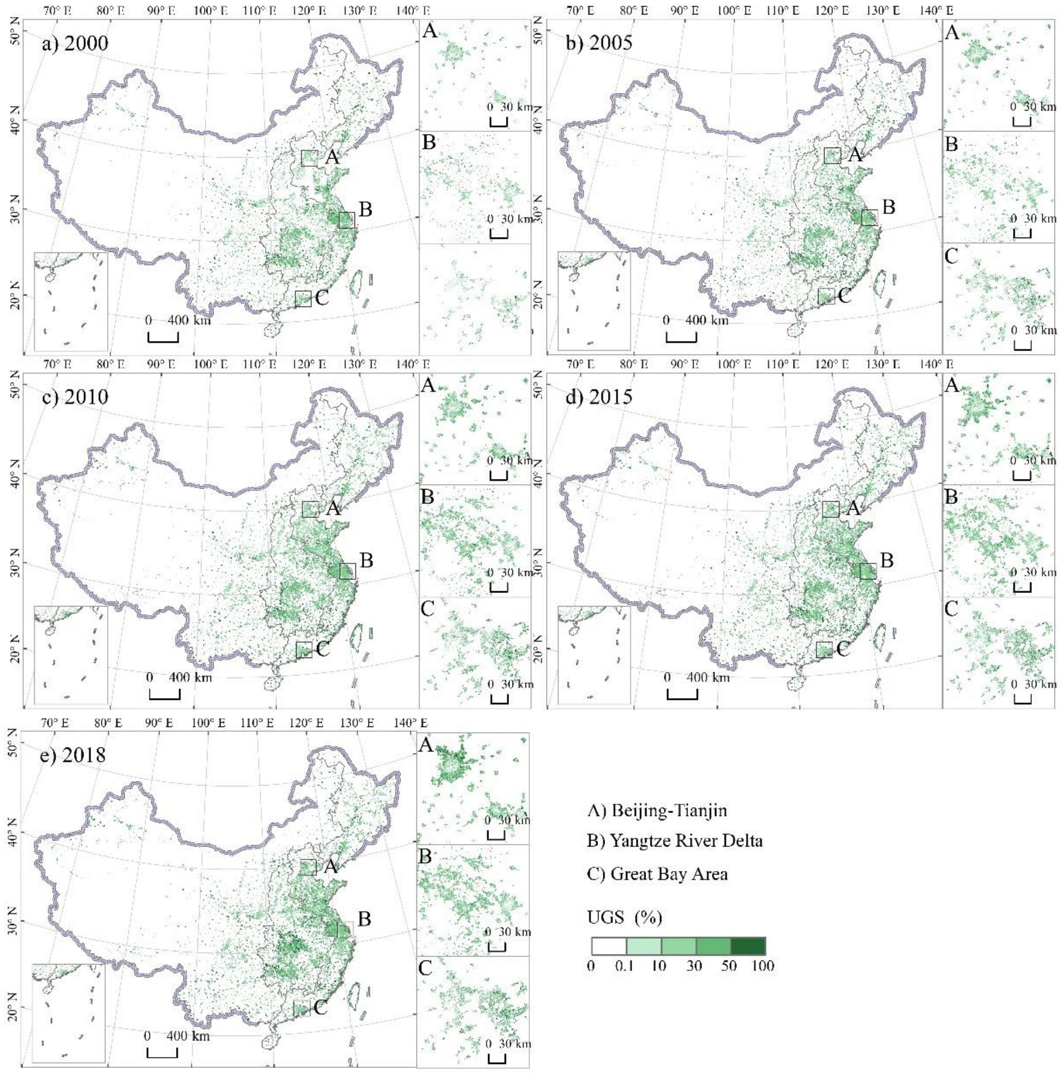

Figure 8: **Spatial distribution of urban green space (UGS) in 2000–2018 across China. (The administrative boundaries were provided by National Geomatics Center of China (http://www.webmap.cn))**

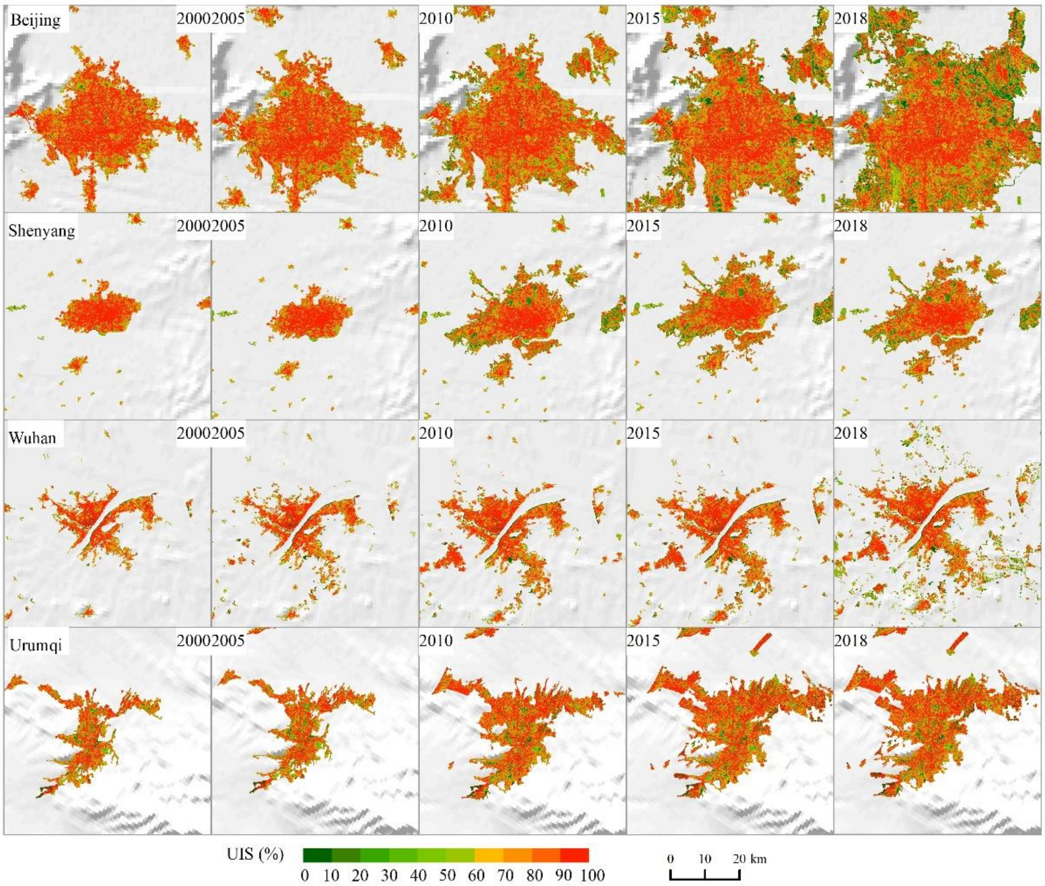

Figure 9: The change of urban impervious surface (UIS) in selected cities from coastal, central, eastern and western zones from 2000 to 2018. (DEM dataset was downloaded from SRTM 90 m Digital Elevation Data (http://srtm.csi.cgiar.org/))

Comment 2: Third, the modeling was based on an existing product (i.e., CLUD), which was based on visual interpretation if I am correct). More details about how urban boundary was extracted and updated should be stated. Without this information, it is hard for readers to know whether urban expansion captured by CLUD was true urbanization or just hand-drawn inconsistency.

Response: Thank you for your comments. In mapping CLUD, the interpretation symbols were built, and the uniform technological flow and classification system were used in human-computer digitalization interpretation environment. Time series of urban boundaries and their expansions have good performance in accuracy and data quality.

At present human-computer interpretation is generally regarded as a the most accurate method in classifying urban land use/cover changes, especially in detecting changing information and extracting vector polygons as whole geo-features.

Changes in manuscript: We added the fourth part “4. Extraction of urban boundaries from CLUD”, including three sub-sections: the classification system and interpretation symbols, land use and dynamic polygon interpretation, and retrieval of multitemporal urban boundaries.

4 Extraction of urban boundaries from CLUD

4.1 The classification system and interpretation symbols

CLUD with 30-m resolution was developed by the Chinese Academy of Sciences and has been updated from 2000 to 2018 every five or three years. This dataset can delineate land use or land cover change associated with human activities, including urbanization at a scale of 1:100,000 (Liu, Liu, Tian et al., 2005; Liu, Liu, Zhuang et al., 2005; Liu et al., 2010). This product adopted a hierarchical classification system covering the first-level six classes and the second-level twenty-five classes. Here the first-level six classes include cropland, woodland, grassland, water body, construction land, and unused land. The detailed description of each class can be found in previous publications (Liu, Liu, Zhuang et al., 2005; Zhang et al., 2014). The construction land was divided into three second-level classes, including urban land, rural settlements, and industrial and mining lands beyond cities. Urban land was defined as a built-up area of the concentrated construction, i.e. buildings, roads, squares, green infrastructure and other lands for providing the living, industrial production, and ecosystem services for the dwellers of cities or towns (Kuang, 2020a). According to the classification system, the interpretation symbols from the second-level classes were built for the false-color composite images as a reference to aid the human-computer interpretation (Fig. 2) (Zhang et al., 2014).

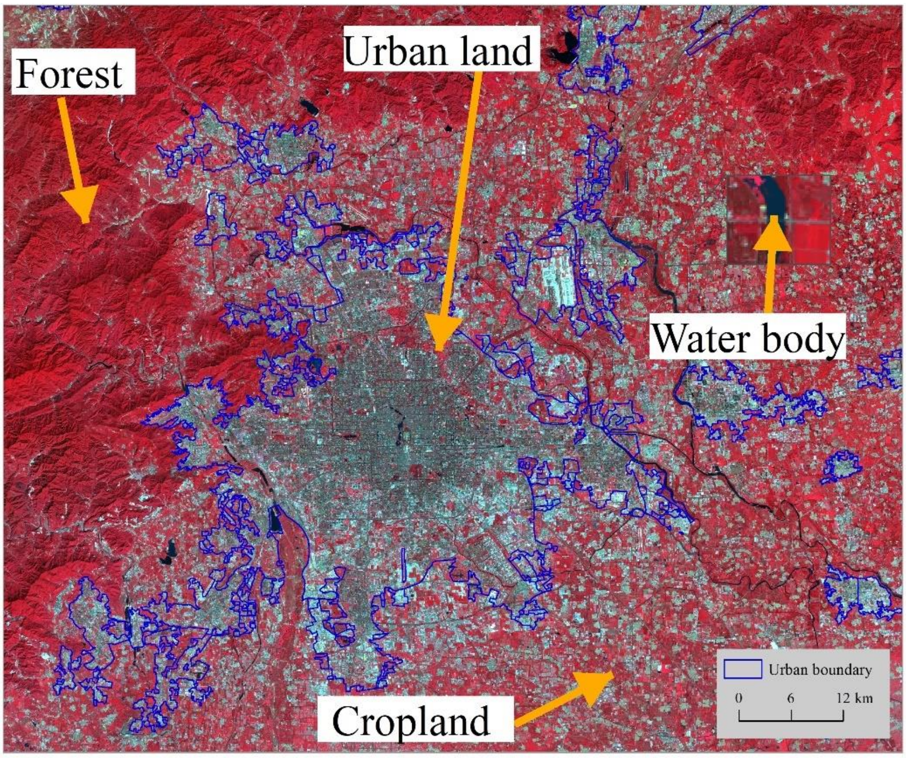

Figure 2: The interpretation symbols and extracted urban boundaries from Landsat images in Beijing city. (The images were provided by Geospatial Data Cloud site, Computer Network Information Center, Chinese Academy of Sciences (http://www.gscloud.cn).

4.2 Land use and dynamic polygon interpretation

According to CLUD classification system, the vector polygons of land use classes in 2000 were digitalized through overlying the false-colour composite images in aid of interpretation symbols and the geoknowledge from each zone (Fig. 3). In the digitalization environment, each vector polygon was assigned with a code of the second-level classes. The vector polygons of land use classes in 2000 were double checked to ensure the correct type in interpretation. The dynamic polygons were extracted through comparing the difference of two-date images and assigned the codes including the types before and after changes (Fig. 3). The land use changes within five or three years were detected using the uniform method. Finally, the land use maps in 2000, 2005, 2010, 2015 and 2018 and their changes at five- or three-year interval were generated for CLUD. The detailed technological flow can be found in previous publications (Liu, Liu, Zhuang et al., 2005; Zhang et al., 2014). An example of land use map in 2010 in Conghua district of Guangzhou city and their dynamic changes in 2010-2015 is illustrated in Fig. 3.

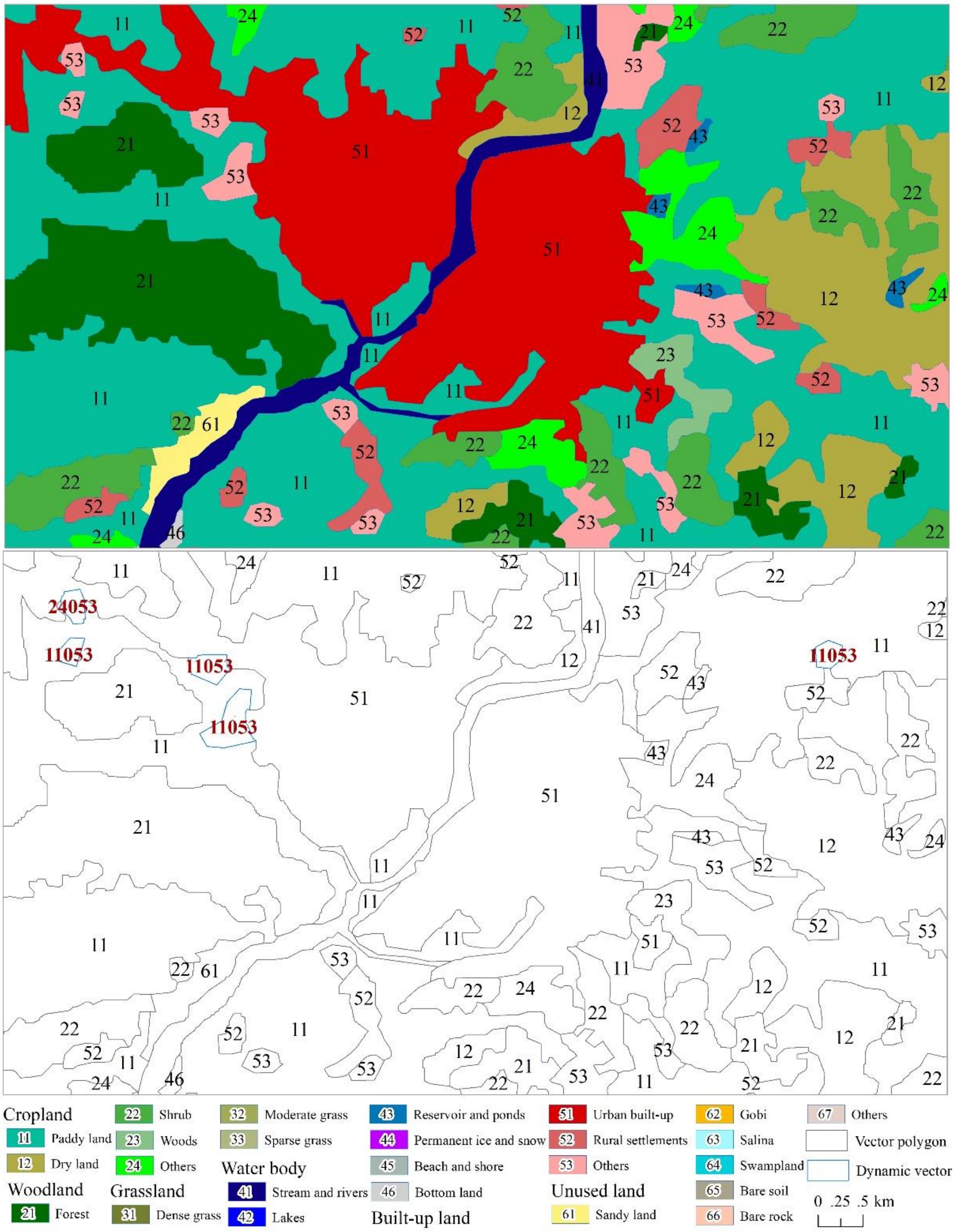

Figure 3: Land use classification and extracted vector polygons as an example with Conghua district of Guangzhou city.

4.3 Retrieval of multitemporal urban boundaries

The vector boundaries of urban extents were extracted from the CLUD land use maps in each period (Kuang et al., 2016). We also examined 10,732 urban vector polygons in 2000.The number of polygons increase to 50,061 in 2018. The vector polygons of urban boundaries were converted to raster data with 30 m x 30 m cell size. The dataset on urban land across China in 2000, 2005, 2010, 2015 and 2018 were generated with 30-m resolution. Here we showed urban boudaries and expansion process with 30-m resolution in cities of Xi'an, Wuhan, Guangzhou and Urumqi (Fig. 4).

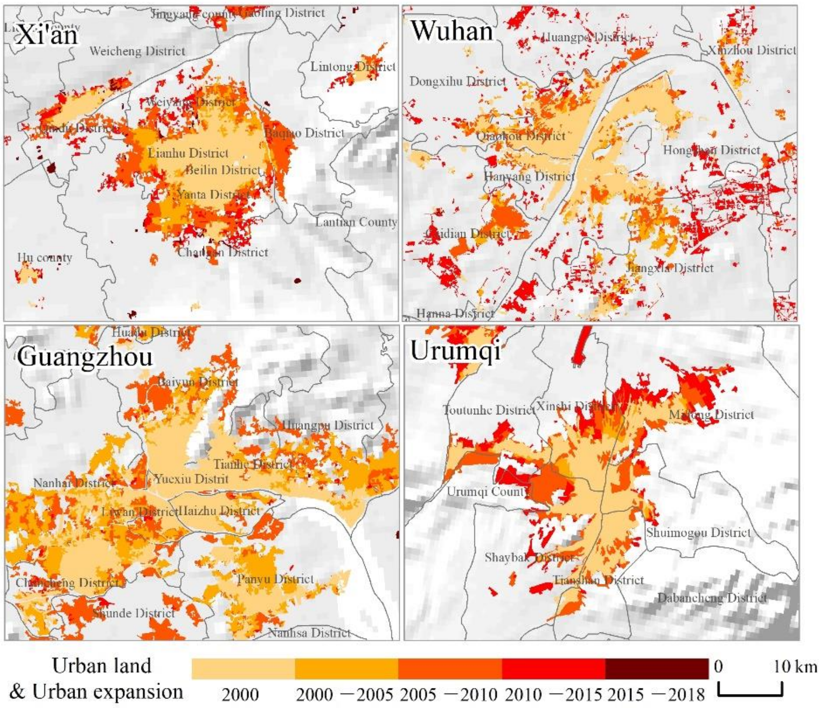

Figure 4: The urban boundaries extracted from CLUD with 30-m resolution in selected cities. (The administrative boundaries were provided by National Geomatics Center of China (http://www.webmap.cn))

Comment 3: How was the accuracy of CLUD assessed? Because the definition of urban in CLUD is more based on administrative perspective instead of surface imperviousness, I want to know more how accuracy of 92-99% was calculated (Lines 149-150).

Response: Thank you for your comments. The accuracy assessment of CLUD in 2015 and 2018 was assessed here, and we also referred a series of publications on CLUD before 2015. The accuracy of the first-level six classes – cropland, forest, grassland, built-up area, water body and unused and of the second-level land use/cove types, including urban land, rural settlements, industrial and traffic lands was assessed using the Google Earth images.

Changes in manuscript: We added the description of the assessment method in "6.1 Accuracy assessment of the CLUD-Urban product".

Comment 4: Last, data uncertainties and limitations should be further addressed. For example, what are spatial and temporal accuracy variations? How consistent was the estimation over time (i.e., is it reliable to use this data set to capture real ISA change)?

Response: Thank you for your suggestions. To address the issue, the land use changes within five or three years were detected using the uniform method. The dynamic polygons were extracted through comparing the difference of two-date images and assigned the codes including the types before and after changes. Therefore, our product has a good consistent in spatial and temporal accuracy.

Changes in manuscript: We added a section "8.3 Limitations of the method and dataset" in discussions in L320-330.