Supplementary Materials

1 Category composition similarity at the global and continental scales

Figure S1 shows the category composition similarity at the global and continental scales. At the global scale, the correlation coefficients of the compositional similarity among the five datasets are quite different, ranging from 69% to 97%, and the inter-annual variation is small.

Globally, the mean value of the correlation coefficient between MCD12 and CCI LC, MCD12 and GLOBCOVER is around 70%, and the mean value between MCD12 and GLCNMO is also under 75%. In contrast, the mean value between CCI LC and the other three datasets is significantly higher than 90% in this comparative analysis. The correlation coefficient of the inter-annual change barely varies. Because of the improved resolution from 1000 m in 2003 to 500 m in 2008 and 2013 for the GLCNMO dataset, the value between CCI LC and GLCNMO increased 4 percentage points in 2008.

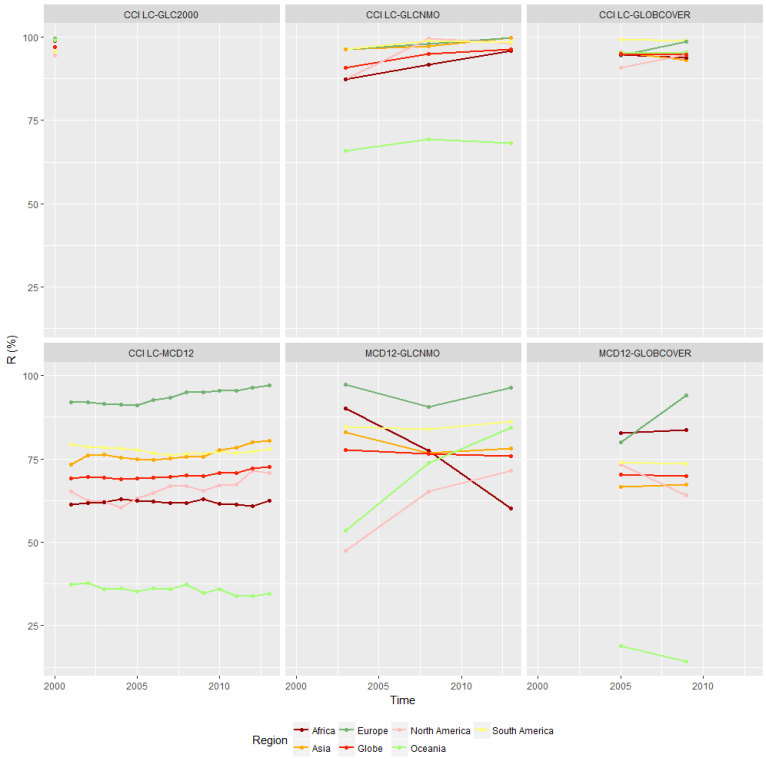

Figure S1. Category composition similarity at the global and continental scales.

As for CCI LC and MCD12, the correlation coefficient (approximately 90%) in Europe was obviously higher than that in other continents. The correlation coefficient in Africa and North America was under 70%, and this value in Asia and South America was 70% - 80%. The correlation coefficient in Oceania was under 40%, whereas the disagreement level was larger than expected. In particular, the correlation coefficient was significantly above 90% between CCI LC and GLOBCOVER, CCI LC and GLC2000. Similarly, the numerical value between CCI LC and GLCNMO was higher than 85%, except in Oceania. However, the numerical value between MCD12 and GLCNMO, MCD12 and GLOBCOVER was irregularly distributed; for

instance, the value between MCD12 and GLOBCOVER in Oceania was under 20%. In summary, the inter-comparison results between MCD12 and GLOBCOVER, MCD12 and CCI LC showed great inconsistencies in Oceania.

The correlation coefficient for the inter-annual change between CCI LC and GLCNMO gradually trended up in Europe, Asia, and South America, and the difference between CCI LC and GLCNMO tended to be slight. From the data of 13-year time series, the correlation coefficient between CCI LC and MCD12 was very stable, and the graphs showed six almost straight lines. However, the correlation coefficient between MCD12 and GLCNMO on three temporal nodes remains unstable, with values plus or minus over 10% in North America, Africa, and Oceania.

2 Overall consistency differences of continental elevations

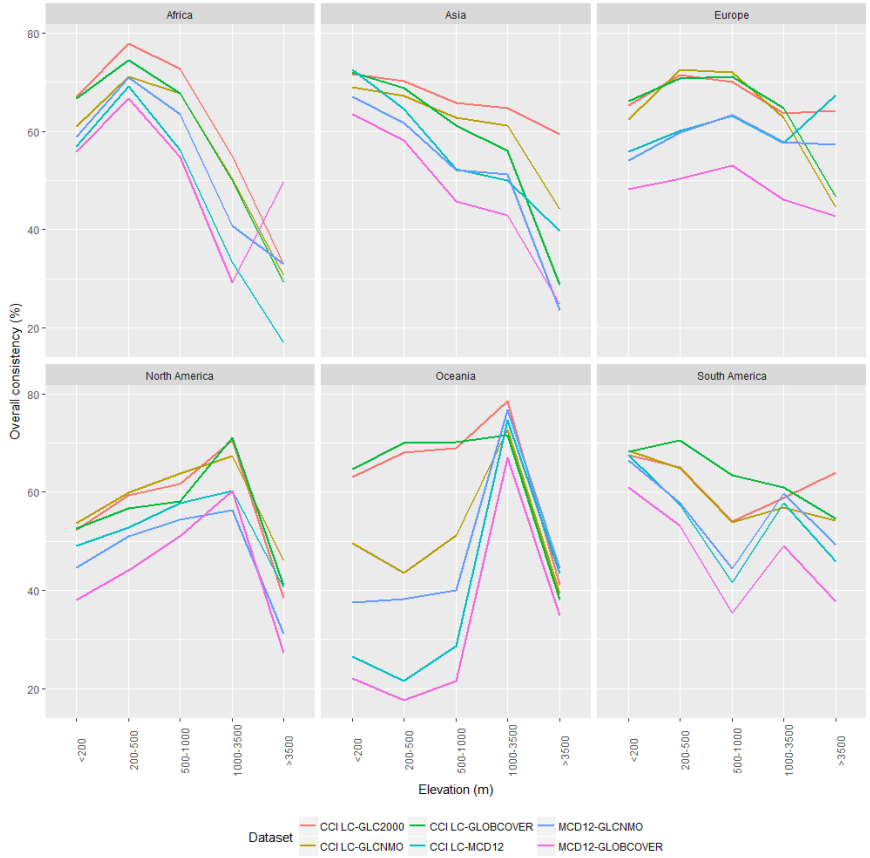

Figure S2. Overall consistency of continental elevations.

Figure S2 illustrates the overall consistency according to the elevation gradient. A total of 24 elevation gradient intervals were present across the six continents. All six datasets exhibited consistent characteristics in two thirds of the intervals. However, at least five datasets had consistent features in 22 intervals, or 91.67% of the total.

For North America, the six datasets exhibited the same characteristics: the overall consistency rapidly increased when elevation was below 3500 m and sharply decreased when the elevation exceeded 3500 m. For Asia, the six datasets exhibited similar characteristics. As the elevation increased, the overall consistency showed a downward trend. The rate of decline

from 1,000 m to 3500 m was relatively flat, and the overall consistency rapidly declined over 3500 m. In Africa, the overall consistency of the five datasets showed similar characteristics with the change of elevation: the characteristics were only somewhat inconsistent above 3500 m. Five datasets showed an upward trend below 500 m and a downward trend above 500 m. Above 3500 m, the comparison of MCD12 and GLOBCOVER showed an upward trend, which was inconsistent with the other datasets. In Europe, the comparison results were inconsistent from 500 m to 1000 m and above 3500 m. In Oceania, the overall consistency was the highest from 1000 to 3500 m, while the overall consistency rapidly decreased above 3500 m. In South America, the four datasets showed similar characteristics: the overall consistency fell below 1000 m, sharply increased between 1000 and 3500 m, and fell again above 3500 m. Notably, the overall consistency (more than 90%) for the comparison results of the datasets showed a downward trend over 3500 m. Although human activities have been greatly reduced above 3500 m, the harsh ecological environment resulted in higher fragmentation of the landscape and the complex interlaced distribution of land objects, which decreased the overall consistency.

3 Category consistency differences of climatic zones

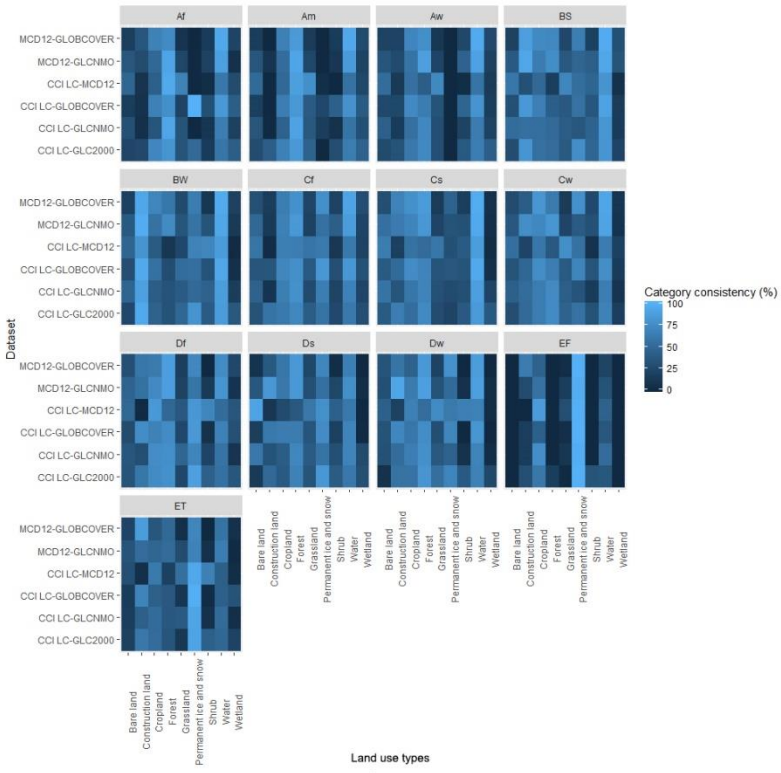

Figure S3. Category consistency of climatic zones.

The comparative results among the six datasets are shown in Figure S3. In the three equatorial climate zone regions (A), namely, Aw, Af, and Am, the category consistency of forest, water and cropland was higher, while the category consistency of other land use types was low. All these climate zones are located in the tropics, so only one dataset identified snow

and ice cover. Af is hot and rainy throughout the year, while Am's dry season is in winter and Aw's dry season lasts more than six months each year. The bare area in Af and Am was far less than that in Aw, so the category consistency was not as good as that of the Aw region.

Compared to the other four large regions, the prominent feature of the arid climatic zone (B) was that the category consistency of bare land and water was very high. In both BW and BS, the category consistency of water exceeded 80%. In the BW area, the category consistency of bare land exceeded 80%, while four of the six comparison results indicated similar situations in BS.

The main feature of B is low rainfall throughout the year, and the climate in most of B is a desert climate. This region includes the bare land in the Sahara Desert and Arabian Peninsula. The area of the bare land is large and concentrated, so the category consistency of bare land was relatively high. The wide distribution of bare land in this area and the difference between the spectral characteristics of water bodies and bare land was large, resulting in a higher category consistency for water.

The prominent feature of the warm temperate climate zone (C) and snow temperate climate zone (D) was that the category consistency of cropland was higher than that of the other regions, with values higher than or close to 60% and even exceeding 80% in some comparison results. C and D contain regions such as India and China's middle and lower reaches of the Yangtze River, which are mostly crop-growing areas. In the polar climatic zones (E), the category consistency of permanent ice and snow cover far exceeded that of other areas because of the continuous and wide distribution of ice and snow.

4 Specifications of spatial consistency on some regional land cover datasets

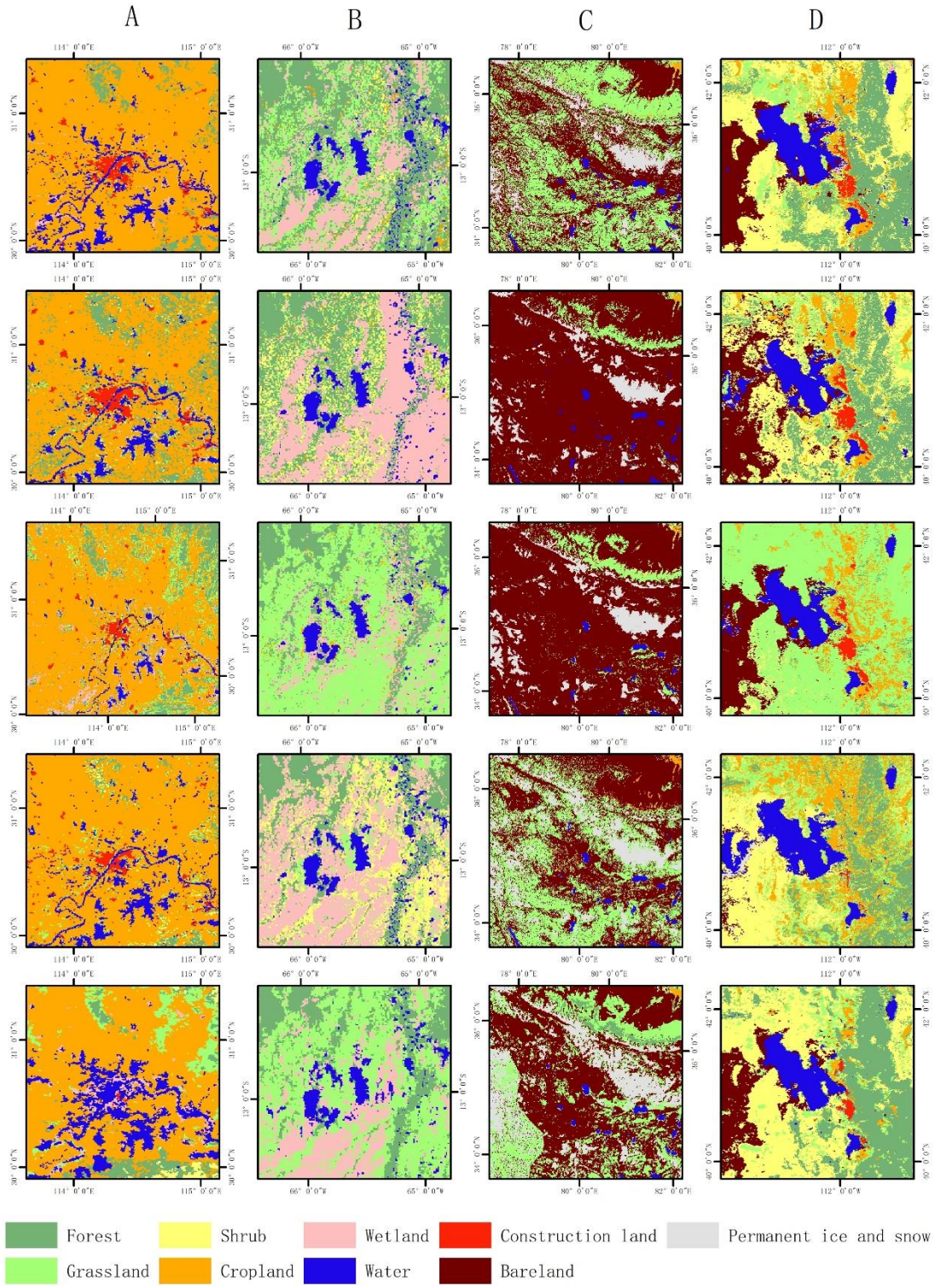

Figure S4. The specifications of spatial consistency on some regional land cover datasets. A, B, C, and D refer to the results of Wuhan of China, Mato Grosso in central Brazil, the Qinghai-Tibet Plateau, and the Utah region of the United States. From top to bottom, they refer to CCI LC (2009), GLCNMO (2008), MCD12 (2009), GLOBCOVER (2009) and GLC2000 (2000).

Figure S4 illustrates the local comparison of Wuhan in China, Mato Grosso in central Brazil, the Qinghai-Tibet Plateau, and the Utah region of the United States and the results of each land-cover dataset are quite different. In group A, each dataset clearly depicted the

contours of the Yangtze River waterways and lakes. However, the GLC2000 classified construction land identified by other datasets into water, while the MCD12 identified relatively small areas of construction land. In group B, MCD12 identified fewer wetlands than other datasets and classified the wetlands identified by other datasets as grassland. All datasets identified the river channel on the right side to a certain extent. In group C, five datasets can identify areas covered by snow and ice, but due to the temporal difference of data sources, the area of snow and ice was inconsistency. There were more inconsistencies in bare land and grassland. In group D, GLOBCOVER classified bare land identified by other datasets as shrub on the left side of the figure and the area of the construction land classified by the GLBOCOVER was significantly smaller than other datasets on the right side of the figure.