# Evaluating the reliability of global historical land use scenarios for forest data in China

YANG Fan[1,2], *HE Fanneng[1], LI Meijiao[1,2], LI Shicheng[3]

1. Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;
2. University of Chinese Academy of Sciences, Beijing 100049, China;
3. School of Public Administration, China University of Geosciences, Wuhan 430074, China

**Abstract:** Global historical land use scenarios are widely used to simulate the climatic and ecological effects of changes in land cover; however, reliability evaluation of these scenarios for data on China's forests is missing. By using a historical document-derived Chinese forest dataset (CHFD) for the years 1700–2000, we evaluated the reliability of data on forests in China over three global scenarios—SAGE (Center for Sustainability and the Global Environment), PJ (Pongratz Julia), and KK10 (Kaplan and Krumhardt 2010)—through trend-related, quantitative, and spatial comparisons. The results show the following: (1) Although the area occupied by forests in China in the SAGE, PJ, KK10, and CHFD datasets decreased over the past 300 years, there were large differences between global scenarios and CHFD. The area occupied by forests in China in the SAGE scenario for 1700–1990 was 20%–40% more than that according to CHFD, and that occupied by forests in the KK10 from 1700 to 1850 was 32%–46% greater than that in CHFD. The difference between the PJ and CHFD was lower than 20% for most years. (2) Large differences were detected at the provincial and grid cell scales, where the PJ scenario was closer to CHFD in terms of total forested area. Provinces with large differences in terms of trend and quantity were 84% and 92% of all provinces, respectively. Grid cells with relative differences greater than 70% accounted for 60%–80% of all grids. (3) These global historical land use scenarios do not accurately reveal the spatiotemporal pattern of Chinese forests due to differences in the data sources, methods of reconstruction, and spatial scales.

**Keywords:** land use; historical forest; natural vegetation; global scenarios; reliability evaluation; China

## 1    Introduction

Changes in land cover caused by land use have significantly altered the Earth in the late Quaternary (Kaplan *et al*., 2017). Long-term land use and land cover change (LUCC) is an important way to understand the processes of global and regional environment changes

(Watson *et al*., 2000; Foley *et al*., 2005; Ellis *et al*., 2013; Zhang *et al*., 2019). Numerous scholars have therefore attempted to reconstruct past LUCC at different spatiotemporal scales (Li *et al*., 2018). Forests, as among the most widespread forms of land cover on Earth (Darby, 1956; Hughes and Thirgood, 1982), have played an important role in regulating climate, the surface radiation budget, and greenhouse gases (House *et al*., 2002; Werf *et al*., 2009). Thus, forest data over long time periods are essential to modeling global or regional climate changes and their ecological effects.

Considerable progress has been made in reconstructing historical forest cover at the local and global scales in recent years (He *et al*., 2019). On the regional scale, scholars have estimated the historical forested areas of several regions based on historical archives. For instance, Clawson (1979) reconstructed the forested area of the United States from 1800 to 1977 using forest survey data. Kaplan *et al*. (2009) built the anthropogenic deforestation scenario of Europe using a sigmoidal relationship between forest cover and historical population density. He *et al*. (2015) created a spatially explicit forest cover for China for 1700–2000 using historical documents, modern surveys, and inventory data. At the global scale, several well-known global historical land use scenarios have featured historical forest data, such as KK10 (Kaplan and Krumhardt, 2010) (Kaplan *et al*., 2011), PJ (Pongratz Julia) (Pongratz *et al*., 2008), and SAGE (Center for Sustainability and the Global Environment) (Ramankutty and Foley, 1999).

Global historical land use scenarios have been widely used to simulate climate change and carbon emissions as a result of anthropogenic activities at both the global and regional scales (Houghton and Hackler, 2003; Brovkin *et al*., 2004; Kaplan *et al*., 2012; He *et al*., 2014). However, there are large differences and uncertainties in these global scenarios due to the use of different methods and proxy data. Uncertainties in these scenarios have propagated into subsequent simulation studies. For example, Kaplan *et al*. (2011) reported that global carbon emissions, based on the KK10 land use scenario, from 8000 years ago to AD 1850 were greater by above 120% than the results calculated by HYDE 3.1 (a global cropland and pasture database covering the period 10,000 BC to AD 2000) (Klein Goldewijk *et al*., 2011). Therefore, what is the uncertainty in global historical land use scenarios, and how should we evaluate these scenarios?

Using local historical documents (Li *et al*., 2019b) or natural evidence (pollen data) (Kaplan *et al*., 2017) to reconstruct historical data helps accurately reflect the spatial and temporal characteristics of regional LUCC. It is, therefore, a significant means of carrying out a regional evaluation of global historical land use scenarios. In China, scholars have evaluated uncertainties in cropland and grassland/pasture in the HDYE, PJ, and SAGE scenarios using globally unique historical document-derived reconstructions, and the results indicate that these scenarios do not capture information on spatial and temporal changes in these land categories (Li *et al*., 2010; He *et al*., 2013; Zhang *et al*., 2013; He *et al*., 2018). However, the reliability of data on forests in China has not yet been evaluated. This study therefore aims to evaluate the uncertainty in the forested area of China in KK10, PJ, and SAGE based on the Chinese historical forest dataset (CHFD) compiled by He *et al*. (2015). The results can provide a useful reference for further understanding and using these global historical land use scenarios.

## 2    Materials and methods

### 2.1    Data sources

Data on forest cover were available in the following global land use scenarios: SAGE, PJ, and KK10. The details of these global scenarios and CHFD are given in Table 1.

**Table 1**    Details of the SAGE, PJ, KK10, and CHFD datasets

| Datasets | Thematic coverage | Temporal coverage | Temporal resolution | Spatial resolution |
|---|---|---|---|---|
| SAGE | Cropland, natural vegetation (forest, grassland) | 1700–1992 | 1–50 a | 0.5°×0.5° |
| PJ | Agricultural area (cropland, pasture), natural vegetation (forest, grassland, shrub, and tundra) | 800–1992 | 1 a | 0.5°×0.5° |
| KK10 | Anthropogenic deforestation | 8000–1850 | 1 a | 5′×5′ |
| CHFD | Forest | 1700–2000 | 5–50 a | 10 km×10 km |

(1) SAGE scenario. This database was first launched in 1999, and mainly contains data on cropland, forests, and grassland (Ramankutty and Foley, 1999). Based on the assumption that the remote sensing-based cropland pattern for 1992 represents the spatial distribution of historical cropland, the authors first synthesized the remote sensing-derived DISCover dataset with historical inventory data to estimate cropland areas from 1700 to 1992. They then overlaid the estimated cropland data over a potential vegetation dataset to obtain historical changes in natural vegetation (including forest data) for the same period. The authors then updated the SAGE cropland (Ramankutty *et al*., 2010) and added pasture data from 1700 to 2007 (Ramankutty, 2012) using population data as a proxy for the area of land use. Considering that the historical forest data in SAGE had not been updated, the data assessed in this study are from the first version, SAGE1999.

(2) PJ scenario. The agricultural areas and land cover of PJ were released in 2008 (Pongratz *et al*., 2008). The authors made regional revisions to the SAGE cropland to create a global cropland for 1700–1992. SAGE pasture data from 1992 were then extended to the period 1700–1992 by combining the rates of change from HYDE (Klein Goldewijk, 2001). The authors subsequently assumed that the ratio of area used per capita for cropland to that used for pasture and the pattern of agricultural areas were constant prior to 1700, and expanded the agricultural areas to AD 800 using population data as a proxy for the agricultural area. Finally, the reconstructed agricultural data were overlaid over a map of potential vegetation to estimate changes in natural vegetation (including forests) from AD 800 to 1992 (https://cera-www.dkrz.de/WDCC/ui/cerasearch/q?query=LAND_COV_scen_1700_1992).

(3) KK10 scenario. Initially, the anthropogenic deforestation scenario for Europe was modeled for the past 3,000 years based on an S-shaped relationship between population density and forest cover on the usable land (Kaplan *et al*., 2009). Using this empirical method, the authors estimated global anthropogenic deforestation between 8 ka BP and AD 1850 using population data as a proxy for anthropogenic deforestation and potential net primary productivity as a limiting factor for this scenario (Kaplan *et al*., 2011) (https://doi.pangaea.de/10.1594/PANGAEA.871369).

(4) Chinese historical forest dataset (CHFD). Based on historical documents, modern

surveys, and statistics as well as previous studies, He *et al*. (2008) estimated the area of provincial forests in China between 1700 and 2000. Subsequently, they devised a forested area allocation model by selecting such limitation factors as altitude, slope, and potential productivity of climate, and divided the provincial forested area into grids at a resolution of 10 km (He *et al*., 2015).

On the whole, maps of forest cover in China for SAGE and PJ were modeled by deducting anthropogenic land use from potential natural vegetation. The deforestation data in KK10 were based on the relationship between the forested area and population data observed in Europe. Current evaluations indicate that cropland and pasture in these global scenarios do not capture the history of land use changes in China; for instance, the area of cropland in the traditional cultivated region of China from SAGE was 112% more than regional estimates (He *et al*., 2013). As a result, significant uncertainties might occur in the forest cover maps generated using these land use data. The CHFD dataset depends on analyzing a large number of historical archives of China and using corresponding methods of reconstruction. It can better capture historical changes in forested areas in China.

## 2.2   Data processing

Owing to the different spatial resolutions and temporal intervals of the SAGE, PJ, KK10, and CHFD datasets, we preprocessed them to facilitate comparison.

(1) Unifying spatial resolutions. The spatial resolutions of SAGE, PJ, and KK10 were 0.5°, 0.5°, and 5′, respectively, while the resolution of CHFD was 10 km (Table 1). To facilitate spatial comparison, CHFD was projected and resampled to 0.5°, and KK10 was up-scaled to 0.5°. At the provincial scale, we adopted the scheme of 25 provinces as in He *et al*. (2015).

(2) Selecting temporal ranges and slices. We selected 1700–1990 as the rang of comparison time of the SAGE, PJ, and CHFD datasets, and 1700–1850 for KK10 given the inconsistency in temporal extent and intervals of the four datasets. Then, 1700 was chosen as the starting point, and the temporal slices were selected at 20-year intervals.

(3) Extracting area occupied by forests in China in KK10. KK10 contains historical deforestation data rather than data on the actual forested area. These data could not be directly compared with those of the other three datasets. Therefore, following the modeling method of KK10, we combined the gridded dataset of land suitability (Ramankutty *et al*., 2002) with that of potential natural vegetation (Ramankutty and Foley, 1999) to create one for potential forest vegetation. Subsequently, the historical forested area of China in this scenario was obtained by deducting deforestation areas from potential forest vegetation.

## 2.3   Evaluation method

According to Section 2.1, the forested area in CHFD was the closest to the fact among the four available datasets. In this study, trend-related, quantitative, and spatial comparisons were used to evaluate the uncertainty in historical data on forests in China in these global scenarios.

(1) Trend-related comparison. The trend-related comparison reveals the consistency and differences among multiple target data items at a macro-scale. The dynamic degree of land use was used here as an indicator of trend-related comparison to characterize trends in areas

occupied by forests in China (Wang and Bao, 1999; Zhu and Li, 2003; Li *et al*., 2019a). The equation for this is as follows:

$$D_f = (F(t_1) - F(t_0)) / F(t_0) \times (1/t) \times 100\% \tag{1}$$

where $D_f$ denotes the dynamic degree of the forested area (%). If $D_f$ is greater than zero, it means that the forested area increases over time; otherwise, the forested area decreases. The higher the absolute value of $D_f$ is, the higher is the intensity of change in the forested area. $F(t_0)$ and $F(t_1)$ refer to the forested areas in years $t_0$ and $t_1$, respectively. Owing to the different intervals among the datasets, we set the unit of $t$ as 10 years, and this index represented changes in forested areas every 10 years in the study period.

(2) Quantitative comparison. This method describes differences in absolute values among datasets. In this study, the logarithmic difference ratio was used to reflect the quantitative differences in forested areas in different datasets, and this is expressed as follows:

$$D_q(t) = \ln(F_g(t) / F_C(t)) \times 100\% \tag{2}$$

where $D_q(t)$ is the natural logarithm of the difference ratio in year $t$, and $F_g(t)$ and $F_C(t)$ denote the forested area in global scenarios and CHFD in year $t$, respectively.

(3) Spatial comparison. Spatial comparison is an analysis at the grid scale that characterizes the variation in the spatial distribution of the same land types in different datasets (He *et al*., 2013; Li *et al*., 2010). We used relative biases to describe differences in the spatial patterns of forests between the global scenarios and the regional results. The equation takes the form:

$$D_r(i,t) = (F_g(i,t) - F_C(i,t)) / F_C(i,t) \times 100\% \tag{3}$$

where $D_r(i,t)$ denotes the relative bias rate of grid $i$ in year $t$ between the global scenarios and CHFD, and $F_g(i,t)$ and $F_C(i,t)$ refer to the forested area in the global scenarios and CHFD in grid $i$ in year $t$, respectively.

## 3  Results and analysis

### 3.1  Overall changes in forested area

The trend of change in forested areas since the year 1700 represented by the datasets reflected a decrease, but estimates for forested areas varied widely between the global scenarios and regional results for China as illustrated in Figure 1.

Forest areas across China in CHFD between 1700 and 1960 underwent a rapid decrease, with a loss of 64% of the area from 1700 and a forested area dynamic degree of –2.45%. The sharpest increase in the forested area was observed during 1960–1998 by 83%, with a forested area dynamic degree of 21.87%. Two stages of changes in forested areas were detected in both SAGE and PJ; a period of linear decrease from 1700 to 1940, with losses of 40% and 46% compared with 1700, and a forested area dynamic de-
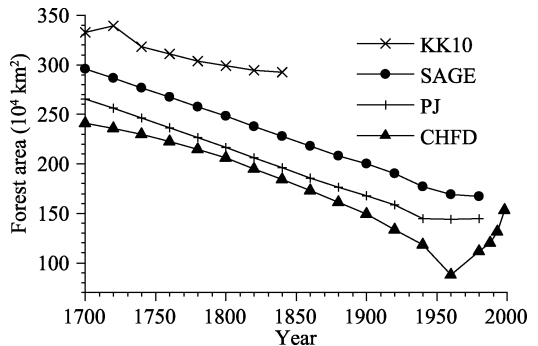


**Figure 1**  Forested areas of China since 1700 according to the SAGE, PJ, KK10, and CHFD datasets

gree of –1.68% and –1.90%, respectively, with a slightly decreased or approximately constant area from 1940 to 1990. KK10 showed a slow decrease from 1700 to 1850, with a total decrease of about 10% and a forested area dynamic degree of –0.67%. The results indicate that changes in forested areas varied significantly between the global scenarios and CHFD.

Large differences in the area occupied by forests in China in the global scenarios and CHFD were observed (Table 2). The area occupied by forests in CHFD was $241.27 \times 10^4$ km$^2$ in 1700, and had decreased to $87.88 \times 10^4$ km$^2$ by 1980, while forested area in SAGE decreased from $296.00 \times 10^4$ km$^2$ to $167.00 \times 10^4$ km$^2$ during this period. The area occupied by forests in China for SAGE was 20%–40% more than that for CHFD; the maximum difference was 65% in 1960. At the same time, the forested area in PJ decreased from $241.27 \times 10^4$ km$^2$ to $143.98 \times 10^4$ km$^2$, closer to that in CHFD than in SAGE. The relative biases in most time slices among them were lower than 20%—only in 1960 did the relative bias reached 49%. From 1700 to 1840, the forested area in KK10 was $292.50 \times 10^4$ km$^2$–$339.63 \times 10^4$ km$^2$, and $184.11 \times 10^4$ km$^2$–$241.27 \times 10^4$ km$^2$ in CHFD. The former was 32%–46% greater than the latter.

**Table 2**  Forest area of China and relative biases among the SAGE, PJ, KK10, and CHFD datasets

| Years | CHFD Forest area ($10^4$ km$^2$) | SAGE | | PJ | | KK10 | |
|---|---|---|---|---|---|---|---|
| | | Forest area ($10^4$ km$^2$) | Relative bias (%) | Forest area ($10^4$ km$^2$) | Relative bias (%) | Forest area ($10^4$ km$^2$) | Relative bias (%) |
| 1700 | 241.27 | 296.00 | 20.44 | 265.23 | 9.47 | 333.04 | 32.23 |
| 1720 | 235.58 | 286.40 | 19.53 | 256.30 | 8.43 | 339.63 | 36.58 |
| 1740 | 229.89 | 276.80 | 18.57 | 246.37 | 6.92 | 318.04 | 32.46 |
| 1760 | 222.81 | 267.20 | 18.17 | 236.41 | 5.92 | 311.20 | 33.41 |
| 1780 | 214.34 | 257.60 | 18.38 | 226.41 | 5.48 | 303.65 | 34.83 |
| 1800 | 205.87 | 248.00 | 18.62 | 216.35 | 4.97 | 299.10 | 37.35 |
| 1820 | 194.99 | 238.00 | 19.93 | 206.22 | 5.60 | 294.55 | 41.25 |
| 1840 | 184.11 | 228.00 | 21.38 | 195.91 | 6.21 | 292.50 | 46.29 |
| 1860 | 172.74 | 218.00 | 23.27 | 185.66 | 7.21 | | |
| 1880 | 160.88 | 208.00 | 25.69 | 176.21 | 9.10 | | |
| 1900 | 149.02 | 200.00 | 29.42 | 167.80 | 11.87 | | |
| 1920 | 133.48 | 190.00 | 35.31 | 158.20 | 16.99 | | |
| 1940 | 117.94 | 177.00 | 40.60 | 144.44 | 20.27 | | |
| 1960 | 87.88 | 169.00 | 65.39 | 143.98 | 49.37 | | |
| 1980 | 111.92 | 167.00 | 40.02 | 144.86 | 25.80 | | |

Overall, SAGE and KK10 overestimated the forested area from 1700 to 1990 across China, and changes in areas occupied by forests varied significantly between them. PJ was closer to CHFD in terms of total forested area.

### 3.2  Provincial forested area

To further reveal the uncertainty in the global scenarios, this study compared them with CHFD at the provincial scale. The comparison in Section 3.1 shows that the total forested area in PJ was the closest to that of CHFD among the global scenarios. Therefore, we carried

out a provincial evaluation of only PJ.

As shown in Figure 2, 13 provinces in PJ exhibited the same trend, as evident across China, featuring a phase of rapid linear decrease followed by a slight decrease. Forested area in 10 provinces decreased more than those in CHFD. For instance, the forested area dynamic degree of Shaanxi Province over this 300-year period in PJ was –1.55%, with a decrease of 43% in terms of area compared with that in 1700. The dynamic degree and decrease in forested area for the same province in CHFD were –0.45% and 13%, respectively. The decreases in forested area in Jing-Jin-Ji (Beijing-Tianjin-Hebei), Jilin, and Chuan-Yu (Sichuan-Chongqing) were smaller than in CHFD. In Jilin, for instance, CHFD had a forested area dynamic degree of –1.82% through the entire period, and the reduction of forested area as recorded by it was 51%, while the dynamic degree and reduction in forested area of this region according to PJ were –1.04% and 29%, respectively.
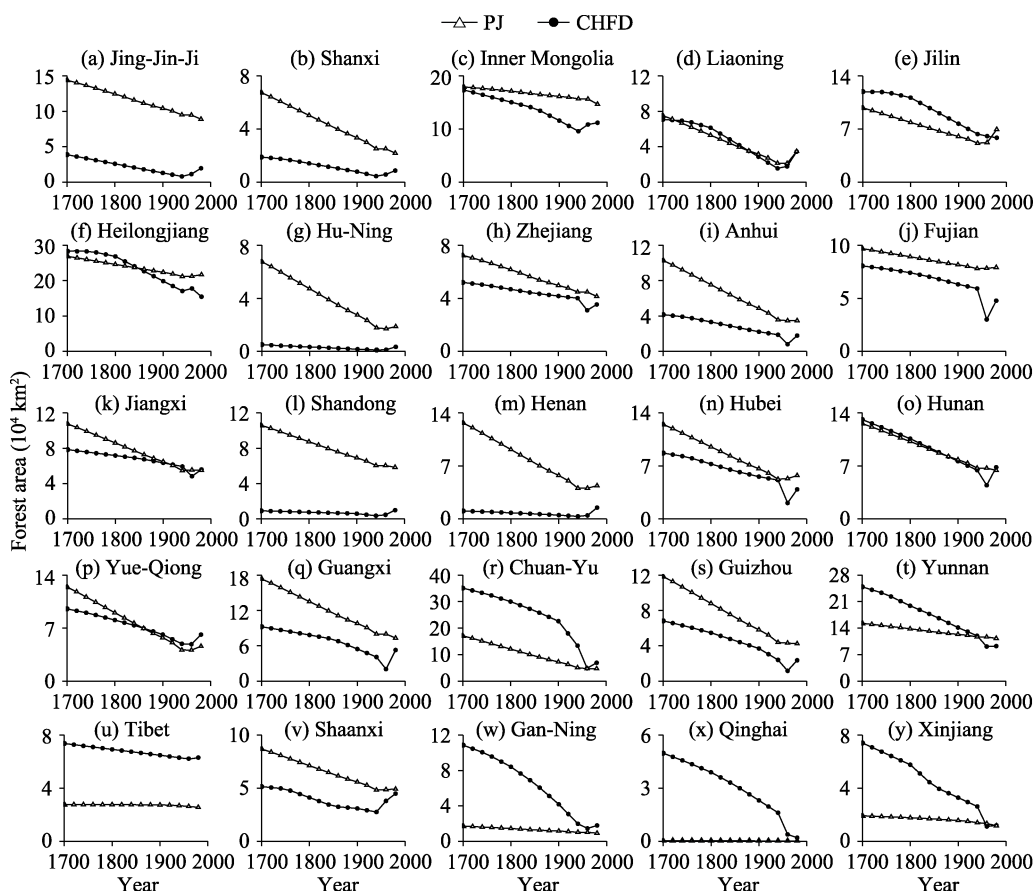


**Figure 2**   Provincial forested area of China for 1700–1990 from the PJ and CHFD datasets

In the PJ scenario, eight provinces exhibited a slow decrease in forested area: Inner Mongolia, Heilongjiang, Fujian, Yunnan, Tibet, Gan-Ning (Gansu-Ningxia), Qinghai, and Xinjiang. The dynamic degrees of their forested areas were between –0.93% and –0.27%. This trend was significantly different from that of CHFD. In the past 300 years, for example, the forested area dynamic degree of Heilongjiang Province in CHFD has been –1.62%, and its forested area has decreased by 45%, whereas the dynamic degree and decline in forested

area for this province in PJ were –0.68% and 19%, respectively. The comparative results suggest that 21 provinces recorded significant variations in forested area, accounting for 84% of all provinces considered.

A total of 23 provinces reflected large differences in area between PJ and CHFD, accounting for 92% of all provinces. The forested areas of 14 provinces in PJ were considerably greater than their counterparts in CHFD. From 1700 to 1980, for instance, the forested area of Hu-Ning (Shanghai-Jiangsu) was between $0.09\times10^4$–$0.50\times10^4$ km$^2$ in CHFD, and $1.70\times10^4$–$6.76\times10^4$ km$^2$ in PJ, a difference as high as 170%–290%. In PJ, seven provinces had forested areas smaller than those in CHFD: Jilin, Chuan-Yu, Yunnan, Tibet, Gan-Ning, Qinghai, and Xinjiang. Throughout this period, for instance, the forested area of Qinghai Province was between $0.20\times10^4$–$4.98\times10^4$ km$^2$ in CHFD, and $0.01\times10^4$–$0.03\times10^4$ km$^2$ in PJ, a difference of 280%–520%. Only the forested areas of Liaoning and Hunan in PJ were closer to those in CHFD, with differences of –15%–22% and –6%–5%, respectively.

The provincial evaluation indicates that 84% and 92% of the provinces had large differences in terms of trend and quantity, respectively, between PJ and CHFD. Therefore, there were significant discrepancies between them at the provincial scale.

### 3.3  Spatial patterns of forest

#### 3.3.1  Comparison of overall spatial pattern

Figure 3 illustrates the overall spatial pattern of forest cover in the PJ and CHFD datasets as well as the relative differences between them. As is shown, the spatial distribution of forests in PJ was in general agreement with that in CHFD, and both showed that eastern China was dominated by forest cover. However, the forest coverage of PJ in the traditional cultivated region of China was significantly greater than that of CHFD. The data in CHFD featured a smaller forest coverage in the North China Plain, the middle lower Yangtze River, and the Loess Plateau than that in PJ because these regions, as traditional cultivated regions of China, had been subjected to high-intensity and large-scale anthropogenic disturbance over long periods. This resulted in large-scale deforestation before the Qing Dynasty. By contrast, these regions in the PJ scenario had a large amount of forest for the entire study period.

In data for the last 300 years, large differences in the range and magnitude of deforestation between PJ and CHFD were found. In CHFD, the forested area decreased significantly in the southwest, northeast, and southeast of China, which accurately captured the spatial and temporal differences in deforestation. For instance, the Qing government had adopted a series of policies to attract immigrants into Sichuan to carry out land reclamation in the late 17th century, leading to a significant loss of local forest cover (He *et al*., 2015). In the mid-19th century, the Qing government abolished its "Prohibit reclamation in Northeast China" policy, and the immigrant population increased exponentially in northeastern China (including Heilongjiang, Jilin, and Liaoning) (Yang *et al*., 2019). As a result, northeastern China became the area with the highest deforestation after southwestern China. By contrast, due to the use of the linear back-scaling reconstruction method, the forested area in PJ decreased at a constant rate, and thus there were no significant spatiotemporal differences in the range and magnitude of deforestation.

In addition, extensive relative biases are shown in Figure 3c. Positive differences were distributed mainly in the North China Plain, the middle lower Yangtze River, and Loess
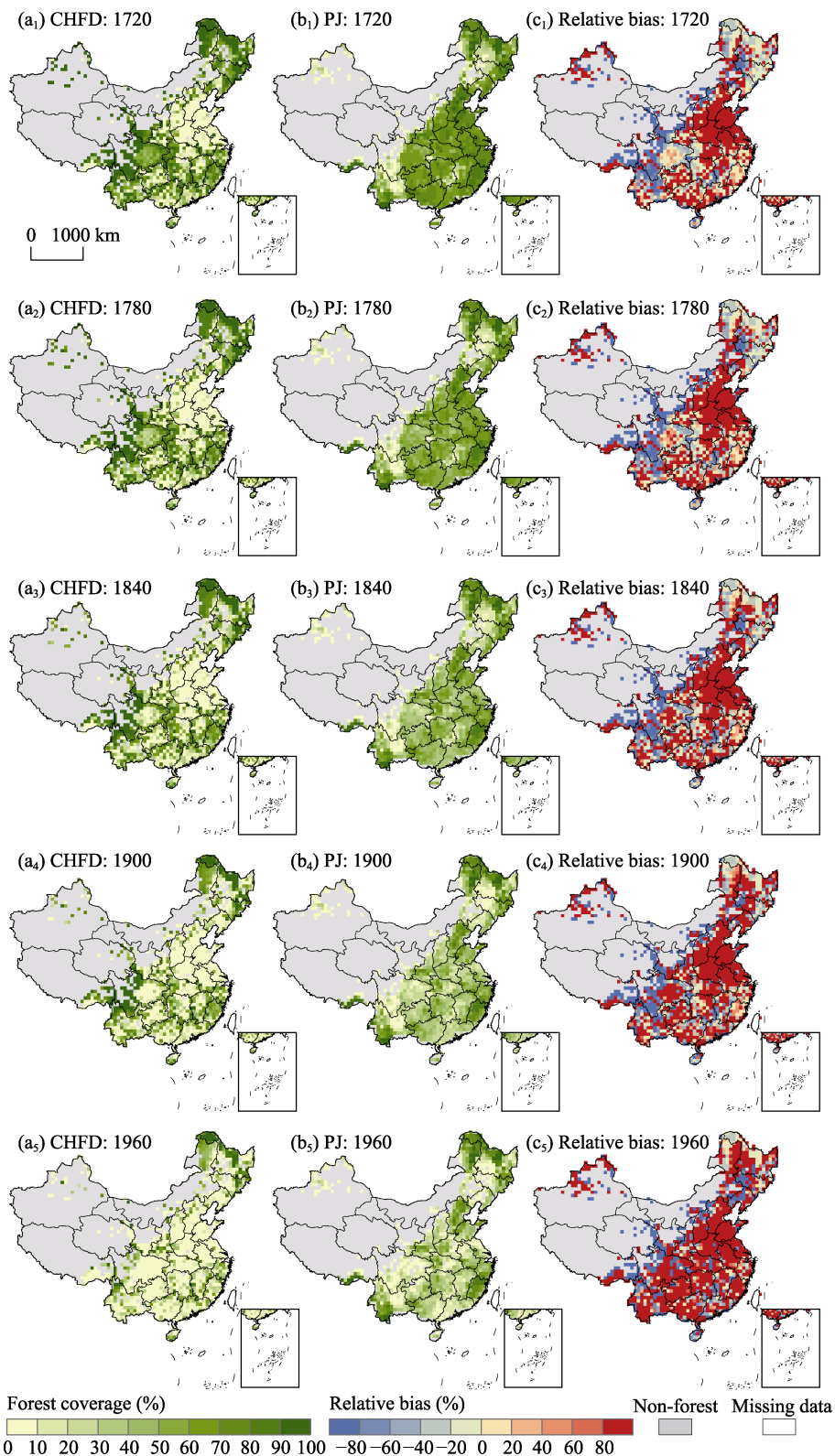
**Figure 3**   The spatial patterns of distribution of forests, and relative biases between the PJ and CHFD datasets

Plateau, and indicated the overestimation of forested areas in PJ for these regions. The positive differences increased over time and were detectable throughout eastern China by 1960. The negative differences in western Sichuan and southeast Tibet reflected the underestimation of forested areas in PJ for these regions. It is clear that the main reason for the extensive relative differences was that deforestation in PJ tended to be represented as uniform.

### 3.3.2  Relative biases on grid cell scale

As shown in Table 3, the percentages of grids with relative biases under 30% (intervals –30%–0% and 0%–30%) were between 10%–26%, and only 3%–9% of grids had relative biases lower than 10%. Grids with relative biases over 70% (intervals >70% and <–70%) accounted for 60%–80% of all grids, and those with relative biases higher than 90% accounted for 55%–75%. The number of grids with large differences increased over time. In 1720, grids with relative biases higher than 90% accounted for 55% of the total, and the proportion of grids with relative biases were lower than 10% accounted for 9%. By 1960, grids with relative biases over 90% exceeded 75%, and those with relative biases under 10% accounted for only 3%. Moreover, significant positive biases were detected between PJ and CHFD. The number of grids with positive biases accounted for more than 80% of all grids while negative biases grids were less than 20%.

**Table 3**  The percentage of grid cells of different relative biases between the PJ and CHFD datasets

| Relative biases (%) | 1720 | 1780 | 1840 | 1900 | 1960 |
|---|---|---|---|---|---|
| <10 | 9.32 | 7.59 | 6.47 | 4.81 | 3.05 |
| 10–30 | 17.14 | 14.51 | 11.73 | 9.47 | 7.18 |
| 30–50 | 7.97 | 9.70 | 9.85 | 8.05 | 5.04 |
| 50–70 | 5.71 | 6.39 | 6.24 | 7.14 | 4.81 |
| 70–90 | 4.66 | 4.81 | 4.74 | 4.21 | 4.66 |
| >90 | 55.19 | 56.99 | 60.98 | 66.32 | 75.27 |

## 4  Discussion and conclusions

Using the historical document-derived CHFD dataset, this study evaluated the reliability of data on forested areas in global historical land use scenarios at the gross, provincial, and grid scales. The main conclusions are as follows:

(1) A gross comparison showed that the reported forested areas varied widely between the global scenarios and CHFD, although the trend was identical in them. SAGE and KK10 overestimated the area occupied by forests in China over the past 300 years; the forested area according to SAGE since 1700 was 20%–40% more than that from CHFD, and that according to KK10 from 1700 to 1850 was 32%–46% more than that in CHFD. The forested area in PJ was closer in value to that in CHFD than SAGE and KK10, with a difference of less than 20%.

(2) The provincial and grid evaluations suggested large differences between the data in PJ and CHFD. Provinces with significant differences in terms of trend and quantity between PJ and CHFD accounted for 84% and 92% of all provinces, respectively. Grids with relative biases over 70% (intervals >70% and <–70%) and 90% (intervals >90% and <–90%) accounted for 60%–80% and 55%–75% of all grids, respectively.

(3) Compared with CHFD, data on forests in China in the global scenarios—SAGE, PJ,

and KK10—were subject to great uncertainty. Differences in aims, data sources, and methods of reconstruction led to significant discrepancies among them. The authors chose remote sensing-derived modern LUCC data available at the global scale as the starting point of the reconstruction, and used the linear back-scaling method to obtain global historical agricultural areas. The reconstructed agricultural data were then overlaid on a map of potential vegetation to estimate the historical forested area. However, at the regional scale, data on forests in China were reconstructed through the analysis and verification of historical documents and survey statistics, and the results accurately reflected historical changes in forested areas in China.

Although our evaluations suggest that land cover data in these global scenarios were uncertain, the method for deducting areas of historical land use from potential vegetation to obtain changes in land cover remains feasible at the global or continental scale, and can be used as a reference for spatially explicit reconstruction in the future, given that the archaeological and paleoecological records capture only land cover at a given site. It is noteworthy that preparing reliable data on potential vegetation and historical land use is essential before using this approach. On the one hand, the potential vegetation data used in global scenarios are mostly based on simulations of the relationship between vegetation and the environment, and the results are uncertain. Using regional historical evidence (e.g., pollen, historical archives, or archaeological investigations) to modify the results of the simulations can contribute to better depictions of natural vegetation. On the other hand, significant progress has been made in reconstructing regional historical land use. The bottom-up approach, i.e., the reconstruction of historical land use from the regional to the global level, is an effective way to improve the reliability of land use data for global scenarios. For example, the Chinese historical cropland data reconstructed based on historical documents have been used in HYDE 3.2. As a result, taking full advantage of historical archives to reconstruct long-term LUCC data is essential for improving the reliability of global scenarios, and meets the needs of regional climatic and ecological simulations.

# References

Brovkin V, Sitch S, Wvon B *et al*., 2004. Role of land cover changes for atmospheric $CO_2$ increase and climate change during the last 150 years. *Global Change Biology*, 10(8): 1253–1266.

Clawson M, 1979. Forests in the long sweep of American history. *Science*, 204(4398): 1168–1174.

Darby H C, 1956. The clearing of the woodland in Europe//Thomas J W L. Man's Role in Changing the Face of the Earth. Chicago: University of Chicago Press.

Ellis E C, Kaplan J O, Fuller D Q *et al*., 2013. Used planet: A global history. *Proceedings of the National Academy of Sciences of the United States of America*, 110(20): 7978–7985.

Foley J A, Defries R, Asner G P *et al*., 2005. Global consequences of land use. *Science*, 309(5734): 570–574.

He F N, Ge Q S, Dai J H *et al*., 2008. Forest change of China in recent 300 years. *Journal of Geographical Sciences*, 18(1): 59–72.

He F N, Li M J, Yang F, 2019. Main progress in historical land use and land cover change in China during the past 70 years. *Journal of Chinese Historical Geography*, 34(4): 5–16. (in Chinese)

He F N, Li S C, Yang F *et al*., 2018. Evaluating the accuracy of Chinese pasture data in global historical land use datasets. *Science China Earth Sciences*, 61(11): 1685–1696.

He F N, Li S C, Zhang X Z *et al*., 2013. Comparisons of cropland area from multiple datasets over the past 300 years in the traditional cultivated region of China. *Journal of Geographical Sciences*, 23(6): 978–990.

He F N, Li S C, Zhang X Z, 2015. A spatially explicit reconstruction of forest cover in China over 1700–2000. *Global and Planetary Change*, 131: 73–81.

He F, Vavrus S J, Kutzbach J E *et al*., 2014. Simulating global and local surface temperature changes due to Hol-

Holocene anthropogenic land cover change. *Geophysical Research Letters*, 41(2): 623–631.

Houghton R A, Hackler J L, 2003. Sources and sinks of carbon from land-use change in China. *Global Biogeochemical Cycles*, 17(2): 1034. doi: 10.1029/2002GB001970.

House J I, Prentice I C, Quéré C L, 2002. Maximum impacts of future reforestation or deforestation on atmospheric $CO_2$. *Global Change Biology*, 8(11): 1047–1052.

Hughes J D, Thirgood J V, 1982. Deforestation, erosion, and forest management in ancient Greece and Rome. *Journal of Forest History*, 26(2): 60–75.

Kaplan J O, Krumhardt K M, Gaillard M-J *et al*., 2017. Constraining the deforestation history of Europe: Evaluation of historical land use scenarios with pollen-based land cover reconstructions. *Land*, 6(4): 91. doi: 10.3390/land6040091.

Kaplan J O, Krumhardt K M, Zimmermann N E, 2009. The prehistoric and preindustrial deforestation of Europe. *Quaternary Science Reviews*, 28(27/28): 3016–3034.

Kaplan J O, Krumhardt K M, Zimmermann N E, 2012. The effects of land use and climate change on the carbon cycle of Europe over the past 500 years. *Global Change Biology*, 18(3): 902–914.

Kaplan J O, Ruddiman W F, Crucifix M C *et al*., 2011. Holocene carbon emissions as a result of anthropogenic land cover change. *The Holocene*, 21(5): 775–791.

Klein Goldewijk K, 2001. Estimating global land use change over the past 300 years: The HYDE Database. *Global Biogeochemical Cycles*, 15(2): 417–433.

Klein Goldewijk K, Beusen A, Van Drecht G *et al*., 2011. The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. *Global Ecology and Biogeography*, 20: 73–86.

Li B B, Fang X Q, Ye Y *et al*., 2010. Accuracy assessment of global historical cropland datasets based on regional reconstructed historical data: A case study in Northeast China. *Science China Earth Sciences*, 53(11): 1689–1699.

Li M J, He F N, Yang F *et al*., 2018. Reconstructing provincial cropland area in eastern China during the early Yuan Dynasty (AD1271–1294). *Journal of Geographical Sciences*, 28(12): 1994–2006.

Li S C, Bing Z, Jin G, 2019a. Spatially explicit mapping of soil conservation service in monetary units due to land use/cover change for the Three Gorges Reservoir area, China. *Remote Sensing*, 11(4): 468. doi: 10.3390/rs11040468.

Li S C, He F N, Zhang X Z *et al*., 2019b. Evaluation of global historical land use scenarios based on regional datasets on the Qinghai–Tibet Area. *Science of the Total Environment*, 657: 1615–1628.

Pongratz J, Reick C, Raddatz T *et al*., 2008. A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles*, 22(3): GB3018. doi: 10.1029/2007GB003153.

Ramankutty N, 2012. Global cropland and pasture data from 1700–2007. Accessed from http://www.geog.mcgill.ca/~nramankutty/Datasets/Datasets.html.

Ramankutty N, Foley J A, 1999. Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles*, 13(4): 997–1027.

Ramankutty N, Foley J A, 2010. ISLSCP II historical croplands cover, 1700–1992. In: Hall F G, Collatz G J, Meeson B *et al*. ISLSCP Initiative II Collection. Tennessee: Oak Ridge National Laboratory Distributed Active Archive Center, 1–20.

Ramankutty N, Foley J A, Norman J *et al*., 2002. The global distribution of cultivable lands: Current patterns and sensitivity to possible climate change. *Global Ecology and Biogeography*, 11(5): 377–392.

Wang X L, Bao Y H, 1999. Study on the methods of land use dynamic change research. *Progress in Geography*, 18(1): 81–87. (in Chinese)

Watson R T N I R, Bolin B *et al*., 2000. Land use, land-use change, and forestry//IPCC Special Report-Summary for Policymakers. IPCC.

Werf G R V D, Morton D C, Defries R S *et al*., 2009. $CO_2$ emissions from forest loss. *Nature Geoscience*, 2(11): 737–738.

Yang F, He F N, Li S C *et al*., 2019. Exploring Spatiotemporal Pattern of Grassland Cover in Western China from 1661 to 1996. *International Journal of Environmental Research and Public Health*, 16(17): 3160. doi: 10.3390/ijerph16173160.

Zhang H Y Gao Z J, Shi M J *et al*., 2019. Study of the effects of land use on hydrochemistry and soil microbial diversity. *Water*, 11(3): 466. doi: 10.3390/w11030466.

Zhang X Z, He F N, Li S C, 2013. Reconstructed cropland in the mid-eleventh century in the traditional agricultural area of China: Implications of comparisons among datasets. *Regional Environmental Change*, 13(5): 969–977.

Zhu H Y, Li X B, 2003. Discussion on the index method of regional land use change. *Acta Geographica Sinica*, 58(5): 643–650. (in Chinese)