The gridded dataset of spring phenology of *Fraxinus chinensis* in China from 1952 to 2007

GE Quansheng, DAI Junhu, WANG Huanjiong

(Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China)

Abstract: The gridded dataset of spring phenology of *Fraxinus chinensis* in China from 1952 to 2007 includes 56 annual files in this time period. This dataset represent the first leaf date of Fraxinus chinensis stored as ARCGIS standard format with a spatial resolution of 1°. This dataset is the result of phenological research in China, and could reflect the spatial-temporal change of spring phenology in China over the past half century. The research article based on this dataset was cited by the fifth IPCC assess report.

Keywords: phenology; *Fraxinus chinensis*; first leaf date; spatio-temporal pattern; climate change

1 Introduction

The gridded dataset of spring phenology of *Fraxinus chinensis* in China from 1952 to 2007 is the result of phenological research of the authors, and is the digital resource reflecting the spatial-temporal change of phenology in China over the past 50 years. We have published several articles about phenological change in China[1]. These results attracted attentions of academic community. A study conducted from similar data set by Ge et al. was also cited by the Intergovernmental Panel on Climate Change (IPCC) fifth assess report[2]. In order to play more important roles in resources and environment sciences, we choose to publish this dataset.

2 Metadata of the SpringPhenologyData_China

The descriptions of the the gridded dataset of spring phenology of *Fraxinus chinensis* in China from 1952 to 2007 (SpringPhenologyData_China) are recorded. These information include the dataset full name, dataset short name, corresponding author, authors, geographical

Table 1 Summary of the SpringPhenologyData_China metadata

Full name of dataset	The gridded dataset of spring phenology of *Fraxinus chinensis* in China from 1952 to 2007		
Short name of dataset	SpringPhenologyData_China		
Corresponding author:	GE Quansheng, geqs@igsnrr.ac.cn		
Authors	GE Quansheng, Institute of Geographic Sciences and Natural Resources Research, CAS, geqs@igsnrr.ac.cn		
	DAI Junhu, Institute of Geographic Sciences and Natural Resources Research, CAS, daijh@igsnrr.ac.cn		
	WANG Huanjiong, Institute of Geographic Sciences and Natural Resources Research, CAS, wanghj.12b@issnrr.ac.cn		
Geographical region	Mainland China, Longitude: from 72° to 136°E; Latitude: from 18°N to 54°N		
Year of the dataset	1952-2007		
Spatial Resolution	1°	Temporal frequency	1 year
Data format	ARCGIS ASCII	Dataset size	636KB
Dataset publisher	Global Change Research Data Publishing and Repository DOI:10.3974/		
Data access and services platform	Global Change Research Data Publishing and Repository, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, http://www.geodoi.ac.cn		
	National Data Sharing Infrastructure of Earth System Sciences of China, http://www.geodata.cn		
Academic editors	LIU Chuang, SHI Ruixiang, WANG Zhenxing, HE Shujin		
Data Sharing Policy	The authors of the dataset agree to publish the data here according to the Article I of Data Sharing Policy of the Global Change Data Publishing and Repository, which states that the dataset can be used freely for research, education, and decision making; any users for commercial uses should get formal permission from IGSNRR/CAS.		

region of the dataset content, year of the dataset, number of the dataset tiles, dataset spatial and temporal resolution, dataset format and size, data publisher, data sharing platform and contact information, technical editors, foundation and the data sharing policy. Table 1 below summarizes the main metadata elements of the SpringPhenologyData_China dataset.

3   Methods

The dataset is derived through the monitoring stations data and the development of phenological model, estimation of parameters and validity of model. All of these models and methodology have been published in Journal of Science China-Earth Sciences[1].

4   Component of dataset

The dataset consists of three parts: headfile.txt, FLDdata.zip and Fraxinus chinensis_distribution.txt[3-5].

(1) headfile.txt. This is a header file. The 1st row "ncols 64" means that the data has 64 columns; the 2nd row means that the data has 36 rows; The 3rd row " xllcorner 72" means that the bottom left corner is at longitude 72°E; the 4th row " yllcorner 18" means that the bottom left corner is at latitude 18° N; the 5th row "cellsize 1" means that the spatial resolution of the data is 1°; the 6th row " NODATA_value -9999" means that the value in area beyond China is -9999. The total volume of this file is 1 KB.

(2) FLDdata.zip. This is a compressed file, which is consisted of totally 56 files (from 1952 to 2007). They are the first leaf dates of *Fraxinus chinensis* in 56 years, which are stored as ARCGIS standard TXT format. The users need to copy the content in headfile.txt to
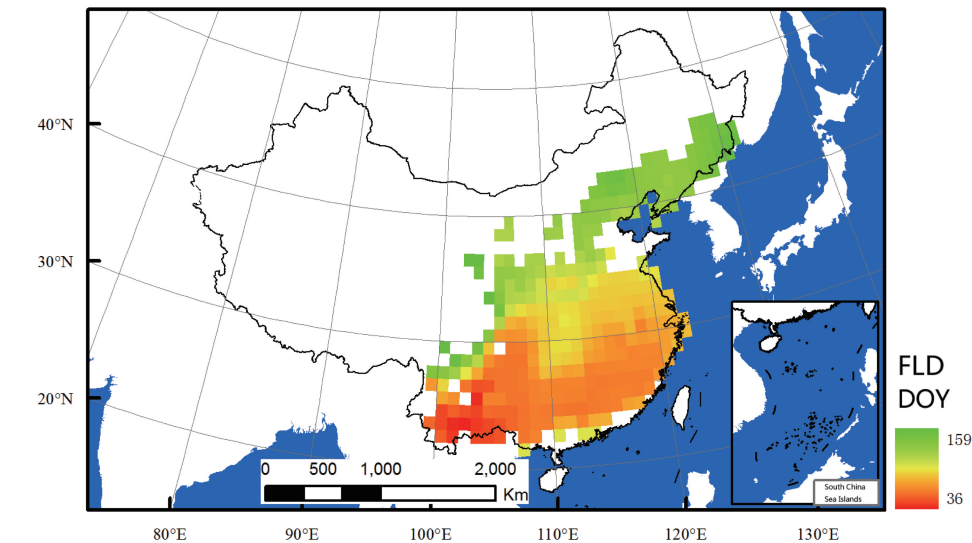
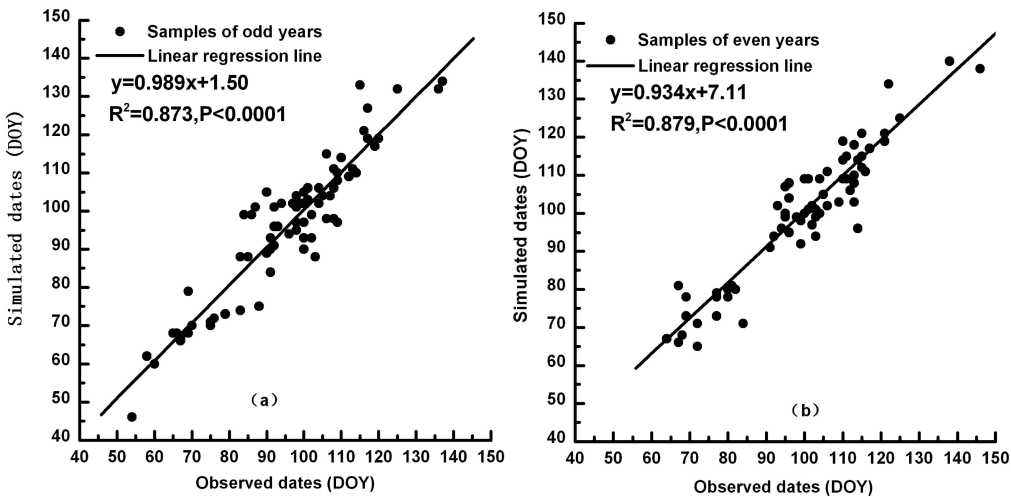Figure 1   The geographical distribution of FLD of *Fraxinus chinensis* in 2007

Figure 2   Internal and external validity of the UniChill model. Percentage variance explained (R^2) and
significance level (P) is shown here. (a) Internal validity; (b) External validity

the first six rows of each file and then converse to ARCGIS GRID format using the "ASCII to Raster" tool in Arcmap. The projection is geographic coordinate system (Spheroid: WGS 84). The total volume of "FLDdata.zip" is 625 KB.

(3) Fraxinus chinensis_distribution.txt. This data represents the distribution range of *Fraxinus chinensis* stored as ARCGIS standard .TXT format. The values 1 and 0 represent that the *Fraxinus chinensis* is distributed in that pixel and not in that pixel, respectively. The value beyond the distribution range is just theoretical value and has no real meanings.

Visualization of the data can be seen in Figure 1.

5   Dataset quality control and validation

The data have been verified by both internal and external validity[6]. Compared with the

74 observed samples of first leaf date (FLD) used to fit the model and the simulated FLD of the corresponding years and sites, their linear regression curve was very close to the $y = x$ baseline (Figure 2a). The explained variance of internal validity was as high as 0.873 ($p < 0.0001$) and the RMSE was 6.1 days. The explained variance of external validity was 0.879 ($p < 0.0001$), and the RMSE was also 6.1 days (Figure 2b), which were calculated between the 70 external samples of FLD and the simulated ones. These results suggested that the data have high accuracy.

6   Conclusion

This dataset is the firstly published data about the annual spring phenology in China at a spatial resolution of $1° × 1°$. Through comparison against the ground- based phenological observation data, this data set has a very high accuracy, and can reflect the impacts of climate changes on biological system together with regional difference of this impact.

References

[1] Wang H, Dai J, Ge Q. The spatiotemporal characteristics of spring phenophase changes of Fraxinus chinensis in China from 1952 to 2007. Science China-Earth Sciences. 2012, 55(6): 991-1000.

[2] Ge Q, Wang H, Dai J. Simulating changes in the leaf unfolding time of 20 plant species in China over the twenty-first century. International Journal of Biometeorology, 2013, DOI: 10.1007/s00484-013-0671-x.

[3] Zhang Qiang. The gridded dataset of surface temperature in China. National Meteorological Information Center, 2007, (http://cdc.cma.gov.cn)

[4] Fang Jingyun, Wang Zhiheng, Tang Zhiyao. Atlas of Woody Plants in China: Distribution and Climate. Beijing: Higher Education Press, 2009.

[5] National Geomatics Center of China, Administrative Boundary GIS Data of China, 1:1 M, 2008.

[6] Chuine I. A unified model for budburst of trees. Journal of Theoretical Biology, 2000, 207: 337-347