



OPEN

DATA DESCRIPTOR

The dataset of main grain land changes in China over 1985–2020

Shidong Liu¹, Li Wang¹ & Jie Zhang²

Continuous, Accurate, and detailed information on main grain land (MGL) areas is crucial for provisioning food security and making policies affecting sustainable agricultural production. It still lacks a long-term MGL distribution dataset with fine spatial resolution. This study aimed to produce a long-term, high-resolution MGL distribution map for China. Here, we developed the change map of MGL with resolution 30 m in China for the period 1985–2020 using the Landsat image-based random forest algorithm on the GEE platform. Finally, the planting intensity, gain time and loss time of MGL was calculated. Results indicate that our mapping results are highly consistent with the annual planting area of various grain crops according to national statistics. A validation based on 3113 field survey samples with a 30-m resolution showed that the overall accuracy of MGL were 93.57%. The full MGL dataset is freely available at <https://doi.org/10.6084/m9.figshare.26212643.v2>.

Background & Summary

Rice, wheat, and maize are the main sources of human food, accounting for approximately 42.5% of the calorie supply and 79.2% of the total cropland¹. However, the rapid increase in global population has led to an increase in food demand, exacerbating the expansion and transformation of cropland use patterns². Meanwhile, climate change, urbanization, and land degradation have led to significant changes in crop cultivation, posing a threat to global food security³. Therefore, accurate and timely information on crop spatial distribution and long-term changes is crucial, especially for agricultural ecosystem management and environmental sustainability policy formulation⁴.

In recent decades, remote sensing technology has been used for mapping crop types for a long time^{5,6}. Due to the fact that different crop types may have similar spectral and texture features during the same growing season in large-scale remote sensing images⁷. Therefore, phenological differences play an important role in large-scale crop classification^{8–10}. Meanwhile, with the development of machine learning technology, the random forest (RF) algorithm has been successfully used for crop classification^{11,12}. The combination of these phenological features and machine learning techniques has facilitated the implementation of large-scale crop classification¹³.

The most important cropland for food security is the main grain land (MGL), mainly plant three food crops: rice, wheat, and maize¹⁴, the three major staple crops that constituted 30.37%, 20.06%, and 40.38%, respectively, of the national total grain production in 2023¹⁵. Meanwhile, the implementation of agricultural measures depends on the MGL use change information. Although there is increasing research on cropland change, the large-scale and long-term MGL changes have not been discussed. The known time series has a resolution of only 500 m and lacks MGL distribution before 2000^{1,16}. Due to the high diversity and complexity of crop planting systems, spatially datasets of MGLs, including single wheat, single rice, single maize, wheat & maize, wheat & rice, double rice, and non-MGL, in China are not available, which hinders research on global food security¹⁷. Meanwhile, many previous studies have focused on mapping single crop types without considering different planting patterns or multiple crop types over long time series¹⁸. Therefore, considering the limitations of short time and low resolution^{19–22}, it is necessary to mapping high-resolution long time series MGL maps. It is urgent to obtain a spatially clear MGL dataset in China in the long term. Accurate and timely MGL datasets are of great significance for relevant scientific research and policy formulation.

To address this issue, we have focused on China, a typical agricultural country, and constructed a MGL dataset by integrating multi-source data and high-resolution satellite imagery. From this, we have quantified the MGL use change. The main aims of this study are: (1) to reveal the crop type and spatial distribution of MGL; (2) to evaluate the accuracy of MGL maps by comparing the ground-truth crop type and grain statistical data

¹Key Laboratory of Remote Sensing and Digital Earth, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, 100094, China. ²Department of Earth System Science Tsinghua University, Tsinghua University, Beijing, 100084, China. e-mail: wangli@radi.ac.cn; zhangj.18s@igsrr.ac.cn

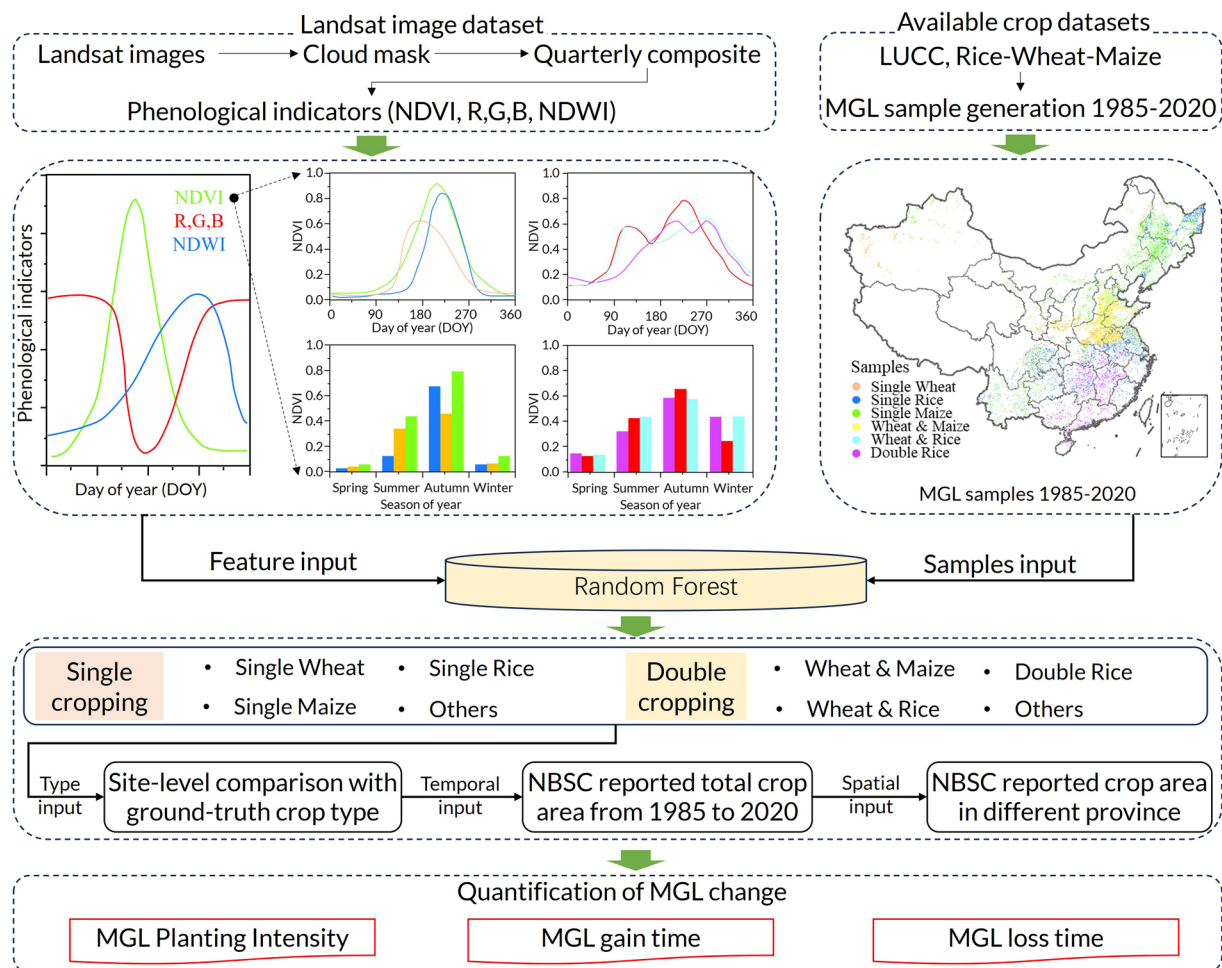


Fig. 1 Study framework.

obtained from the National Bureau Statistical of China (NBSC); (3) to investigate the spatiotemporal patterns of MGL change. This study provides scientific support for optimizing MGL farming practices, ensuring global food security, and promoting agricultural management.

Methods

Study framework. The study framework used for mapping MGL includes data preprocessing, MGL mapping methods, accuracy evaluation, and quantification of MGL change patterns. We first constructed phenological indicators based on the Google Earth Engine (GEE) platform, including Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI). NDVI is often used to reflect surface vegetation characteristics²³, evaluate crop growth status, and is one of the important indicators for identifying crops²⁴. NDWI is a key indicator for distinguishing between water and land, often used to identify rice planting areas^{25,26}. Therefore, phenological change information characterized by NDVI, NDWI, and RGB bands of Landsat are combined to train an RF model for identifying MGL. In addition, we collected all available land use/cover change (LUCC) data and spatial distribution datasets of the major grain lands from 1985 to 2020. Used to construct annual MGL training samples. Combining phenological indicators with MGL samples, a machine learning algorithm based on RF is used to train and identify different MGLs. We developed an MGL dataset, analyzed spatiotemporal changes in cropland and MGL in China over 36 years (Fig. 1).

To ensure the accuracy of MGL mapping results, we adopted three different data validation methods. Firstly, based on field observation data from different years, the accuracy of identifying types is verified at the plot scale. Secondly, by calculating the total MGL area from 1980 to 2020 and comparing it with the data reported by National Bureau of Statistics of China (NBSC), the accuracy of MGL identification at the time series scale is verified. Finally, we validated the spatial accuracy of the dataset based on different crop planting areas reported by NBSC in different provinces. Finally, based on accurate MGL mapping results, we extracted critical crop planting patterns, including crop types, MGL gain time (the time when the non-MGL pixels first transformed into MGL), planting intensity (the total number of years used for planting MGL), and MGL loss time (the time when MGL pixels first transform into non-MGL pixels), using time-series analysis. These indicators are the core indicators reflecting the changes in China's MGL over the past 36 years.

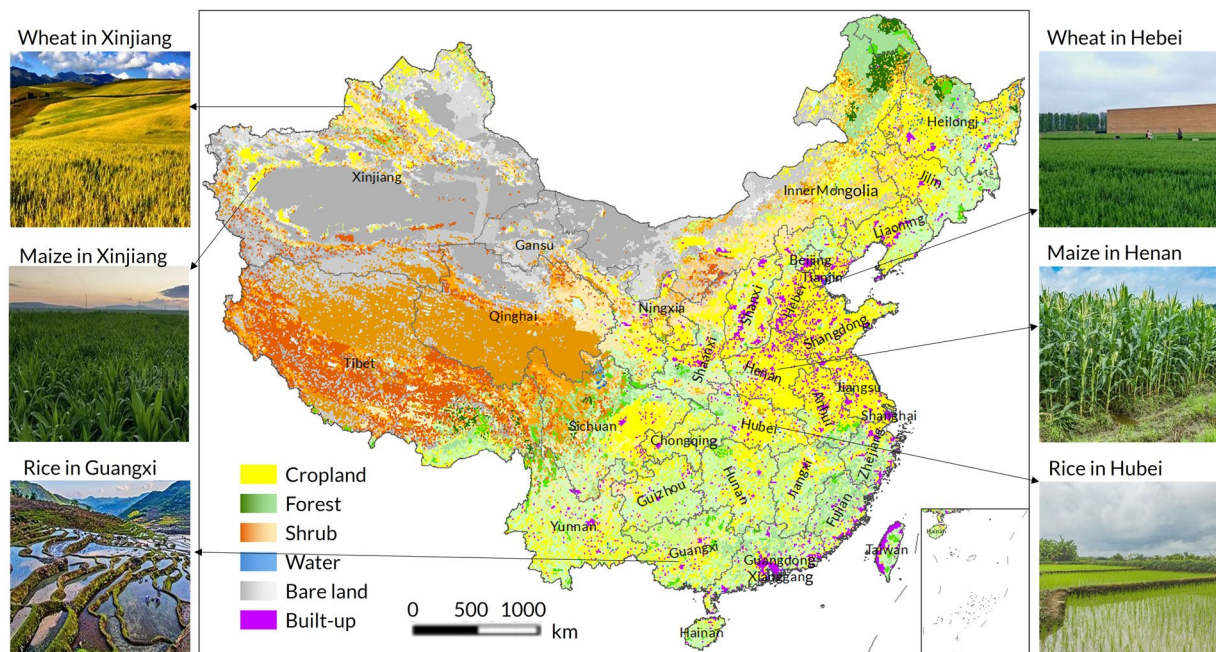


Fig. 2 Study area. The land cover data is sourced from China Land Cover Dataset 1985–2020³².

Data type	Data Name	Resolution	Data sources
Land/cropland Cover	Wheat in China 2001–2020	30 m	http://www.nesdc.org.cn/
	Maize in China 1985–2020	30 m	https://doi.org/10.6084/m9.figshare.16437054
	Rice in China 1990–2016	30 m	https://doi.org/10.57760/sciencedb.15865
	Double rice in China 2016–2020	10 m	https://doi.org/10.12199/nesdc.ecodb.rs.2022.012
	China rice, wheat, and maize planting dataset 2000–2019	1 km	https://doi.org/10.17632/jbs44b2hrk.2
	Annual cropland dataset of China 1986–2021	30 m	https://doi.org/10.5281/zenodo.7936885
	China Land Cover Dataset 1985–2020	30 m	https://zenodo.org/record/8176941
Image data	Landsat images	30 m	Google Earth Engine
Field data	Agro-meteorological data	—	http://data.cma.cn/en
Vector data	Provincial administrative boundary data in China	—	https://www.resdc.cn/
Statistical data	Crop planting area	—	http://www.stats.gov.cn/english

Table 1. Data source of this study.

Study area and data source. This study focuses on China, where the agricultural planting system is highly diverse and complex²⁷ (Fig. 2). Wheat, rice, and maize are the most important staple crops in China²⁸. Globally, wheat, rice, and maize account for 79% of the total grain harvest area²⁹. In fact, in China, these three main crops accounted for about 97% of the country's grain area in 2020 (www.stats.gov.cn/english/).

The study mainly utilized datasets of cropland and MGL distribution, field crop monitoring data, and statistical data (Table 1). The spatial distribution of rice^{19,20}, wheat²¹, and maize²² in China from 1985 to 2020 were collected to assist in producing national MGL training samples³⁰. We employed the Landsat dataset in GEE to generate MGL. Additionally, the annual 30 m cropland dataset in China from 1986 to 2021 was used to constrain the scope of MGL³¹, with the cropland in 1985 being supplemented using the cropland from the China land use dataset in 1985³².

Identification of MGL. We have collected spatial distribution datasets of rice, wheat, and maize in China within a resolution range of 30 m–1 km, spanning from 1985 to 2020. Based on these existing datasets, we have utilized data fusion techniques to create a nationwide fundamental sample dataset of MGL. In the specific process of data fusion, we first expand the spatial range of all available datasets to the entire China. During the expansion process, fill the missing pixels with NoData and resample them to a resolution of 30 m for subsequent statistics and fusion. Secondly, we used a probability distribution model (Eq. (2)) to calculate the probability ($P(T)$) of each pixel belonging to non-MGL (code: 0, N), wheat (code: 1, W), maize (code: 2, M), single rice (code: 3, R), and double rice (code: 4, DR) based on all existing crop datasets (A).

$$A = \{T \mid 0, 1, 2, 3, 4\} \quad (1)$$

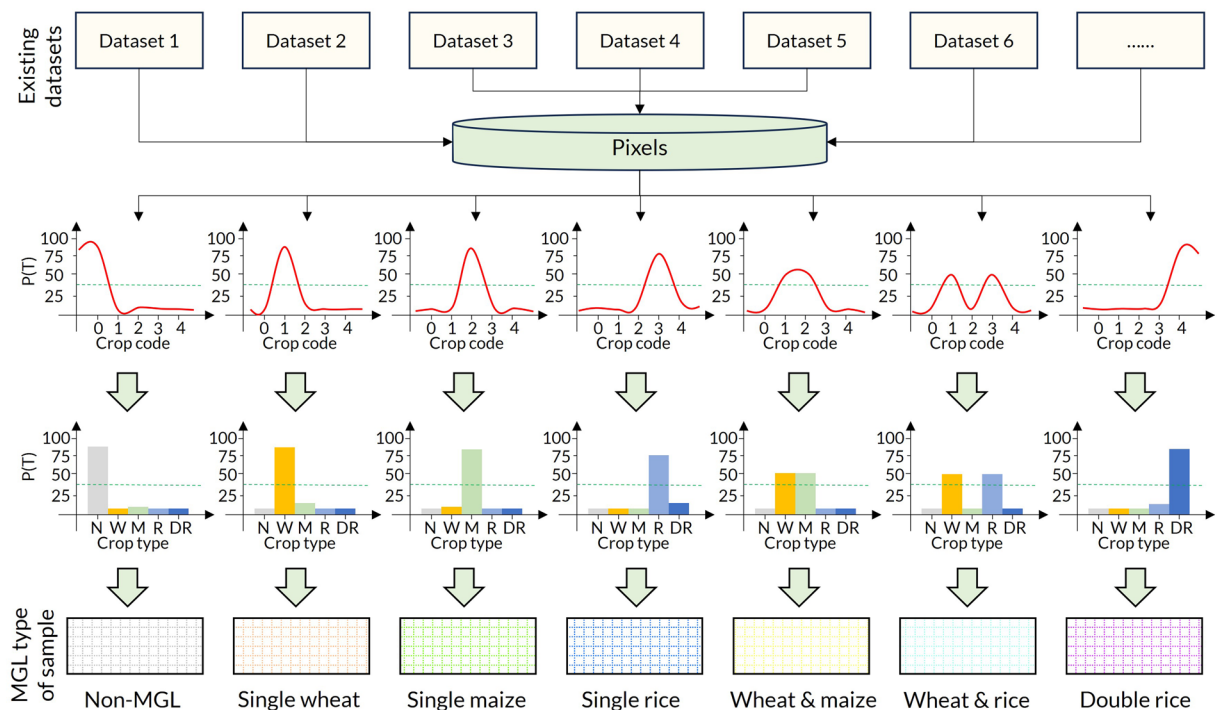


Fig. 3 Data fusion process of the fundamental sample dataset of MGL.

$$P(T) = \frac{N_{Type}}{N_A} \times 100\% \quad (2)$$

Considering that the multiple cropping index of cropland in China does not exceed 3 times, this study assumes that a pixel can belong to no more than 3 crop types in a year, which means that the threshold of the probability distribution model is 33%. Based on the probability distribution characteristics of each crop in each pixel, this study determines the actual type of pixel as the crop with a probability greater than the threshold. If a pixel belongs only to a single crop type other than double rice, it is determined as a single planting MGL type, such as single wheat, single rice, or single maize. If a pixel belongs only to double rice, then the pixel is defined as double rice. If a pixel belongs to multiple crop types, it is defined as an MGL type that has been planted multiple times, such as wheat & maize, wheat & rice, maize & rice. Considering that there are no pixels belonging to both maize & rice, seven basic MGL types including non-MGL were formed in the basic sample dataset (Fig. 3).

Considering 7 types comprehensively, including single wheat, single rice, single maize, wheat & maize, wheat & rice, double rice, and non-MGL, we have automatically selected 100,000 sample points nationwide from the fundamental sample dataset using stratified sampling. Meanwhile, utilizing Landsat imagery from the GEE platform, we have constructed a remote sensing imagery feature set for each year. Available spectral features are synthesized every two months, including red, green, blue, and near-infrared bands, as well as NDVI, NDBI, and NDWI. By combining six spectral features from each year, we have formed a phenological feature dataset. Finally, using the random forest algorithm, we individually mapped the MGL for each year. By merging all types of MGL, the cropland map of China from 1985 to 2020 can be obtained.

Quantification of cropland planting patterns. Based on the MGL dataset from 1985 to 2020, we assigned MGL and non-MGL values of 1 and 0 to all cropland, respectively. The planting intensity gain time and loss time of cropland and MGL was calculated.

Detection of MGL planting intensity. Using time series analysis techniques, we counted the number of times each pixel (C_{ij}) belonged to cropland or MGL over the past 36 years (Eq. (3)), which were named as the cropland planting intensity ($PI_{cropland}$) and the MGL planting intensity (PI_{MGL}), respectively. The attribute range of the PI_{MGL} map is 0–36, where 0 represents non-MGL cropland. The values from 1 to 36 represent the number of times that pixel has been used for PIMGL over the past 36 years. Similarly, In the $PI_{cropland}$ map, 0 represents non-cropland, and the meaning of attribute values is similar to PI_{MGL} .

$$PI = \sum_{k=1985}^{2020} C_{ijk} \quad (3)$$

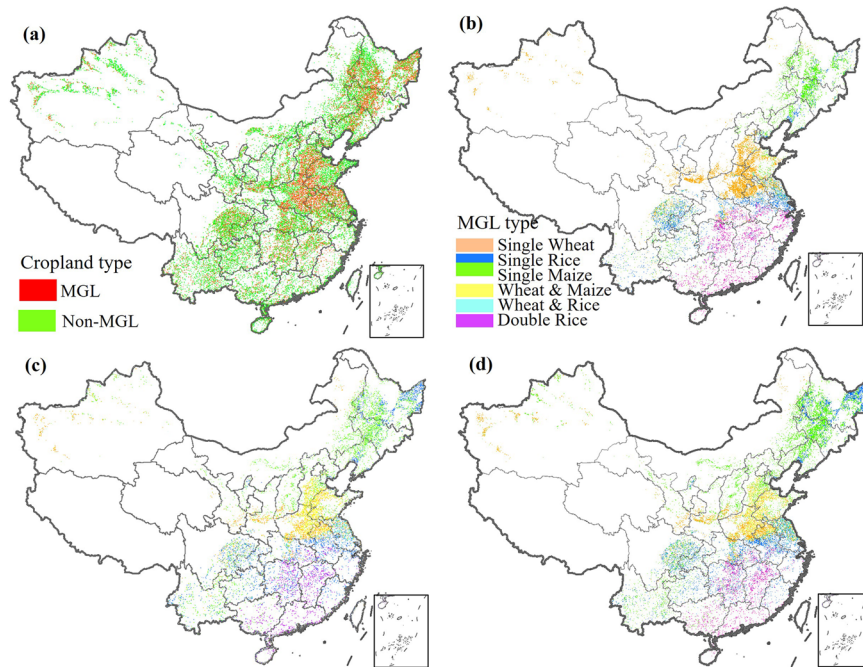


Fig. 4 MGL in China from 1985 to 2020. (a) is the overall distribution of MGL in 2020. (b) is the mapping results of MGL in 1985. (c) is the mapping results of MGL in 2005. (d) is the mapping MGL results for 2020.

Detection of MGL gain time. Using time series analysis techniques, we counted the first time of each pixel (C_{ij}) belonged to cropland or MGL over the past 36 years (Eq. (4)), which were named as the cropland gain time ($GT_{cropland}$) and the MGL gain time (GT_{MGL}), respectively. The attribute range of GT_{MGL} map is 1986–2020, and the values of each pixel represent the time when the pixel first transformed into MGL. The attribute range of the $GT_{cropland}$ map is 1986–2020, and the values of each pixel represent the time when the pixel was first developed into cropland.

$$GT = \min_{k=1985 \rightarrow 2020} [k \times C_{ijk} | C_{ijk} = 1] \quad (4)$$

Detection of MGL loss time. To ensure the stability of the cropland loss time ($LT_{cropland}$) and MGL loss time (LT_{MGL}), we define the pixels (C_{ij}) that have been non-cropland or non-MGL for three consecutive years as stable loss pixels. Using time series analysis techniques, we counted the first time of each pixel (C_{ij}) belonged to $LT_{cropland}$ or LT_{MGL} over the past 36 years (Eq. (5)), respectively. The attribute range of LT_{MGL} map is 1987–2020, and the values of each pixel represent the time when the pixel first transformed into non-MGL. The attribute range of the $LT_{cropland}$ map is 1987–2020, and the values of each pixel represent the time when the pixel was first abandoned or converted into cropland, including long-term fallow. The time for re-cultivation after fallow is recorded on the GT map.

$$V_{ijk} = \max_{k=1985 \rightarrow 2018} (C_{ijk}, C_{ijk+1}, C_{ijk+2}) \quad (5)$$

$$LT = \min_{k=1987 \rightarrow 2020} [k \times (1 - V_{ijk}) | V_{ijk} = 0] \quad (6)$$

Crop planting area statistics. The annual sowing area and the area planted with MGL are different concepts. Due to the fact that the national statistical agricultural planting area is only the total sowing area of a single crop, without the composite MGL area, we cannot directly verify the accuracy of MGL. To verify the accuracy of MGL's mapping area, it is necessary to calculate the annual total planting area for each crop type (A_i), which is the sum of the areas of any MGL type planted with that crop ($AMGL_{ij}$). The existence of the multiple cropping index allows this study to calculate the crop sowing area (Eq. (7)), ensuring comparability with the national statistical agricultural planting area. Please note that we have calculated the double cropping rice area twice when calculating the rice planting area ($m_{ij} = 2$).

$$A_i = \sum_{j=1}^n m_{ij} \times AMGL_{ij} \quad (7)$$

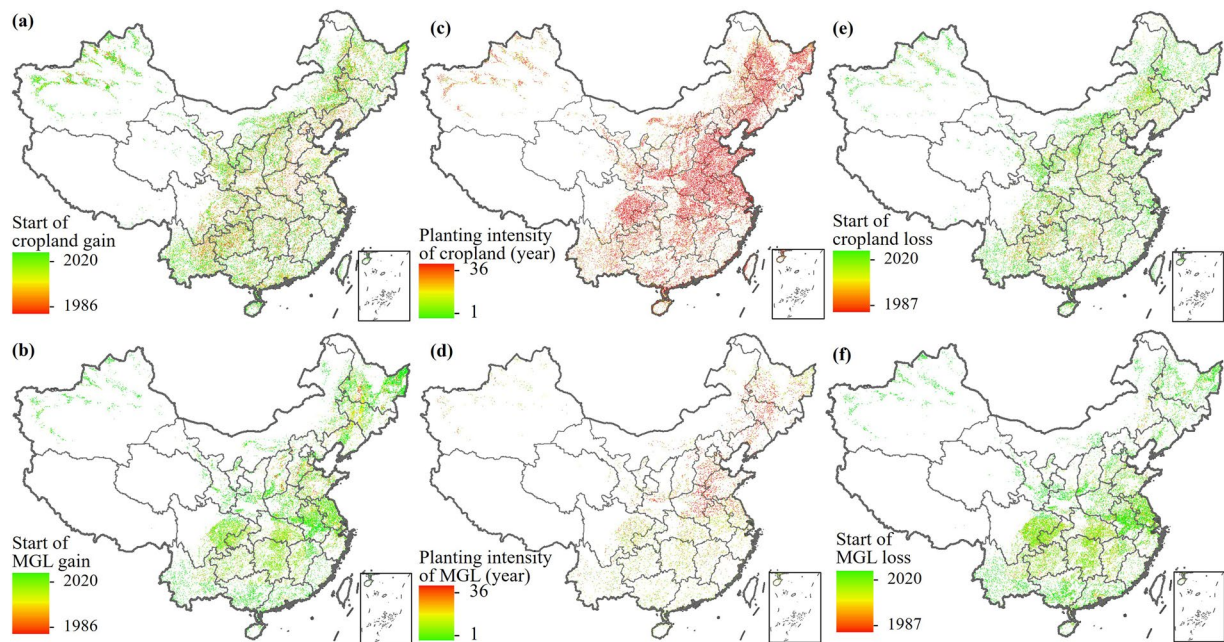


Fig. 5 The change pattern of MGL.

Number of samples		Actual label					
		Total	Rice	Wheat	Maize	Others	Producer accuracy (%)
Estimated results	Total	11345	3360	1750	3122	3113	
	Rice	3284	3119	12	16	137	94.98
	Wheat	1837	25	1680	50	82	91.45
	Maize	3137	76	30	2977	54	94.90
	Others	3087	140	28	79	2840	92.00
	User accuracy (%)	92.84	96.00	95.36	91.23		
	Overall accuracy (%)	93.57					

Table 2. Accuracy assessment using field sites. Note: The ‘Others’ type refers to non-MGL cropland.

Data Records

MGL map in China. The datasets are available at the figshare repository in a Geotiff format with the ESPG: 4326 (WGS_1984) spatial reference system³³. The main grain land changes in China are provided during 1985–2020 (Fig. 4). The dataset includes 36 bands. Bands 0 to 35 represent the MGL maps from 1985 to 2020, respectively. For an annual year MGL map, the code of MGL map ranges from 0 to 7, representing non-cropland, single wheat, single rice, single maize, wheat & maize, wheat & rice, double rice, and non-MGL respectively. Similarly, the cropland dataset can be obtained by simply merging the MGL dataset.

MGL change pattern map. The dataset also includes maps of three map layers of MGL planting patterns, including MGL planting intensity, MGL gain time, and MGL loss time (Fig. 5). The attribute range of MGL gain time map is 1986–2020, and the values of each pixel represent the time when the pixel first transformed into MGL. The attribute range of the MGL planting intensity map is 0–36, where 0 represents non-MGL cropland. The values from 1 to 36 represent the number of times that pixel has been used for PIMGL over the past 36 years. The attribute range of MGL loss time map is 1987–2020, and the values of each pixel represent the time when the pixel first transformed into non-MGL. The planting pattern map of the cropland dataset corresponding to the MGL year has the same pixel attributes as the MGL planting pattern.

Technical Validation

Comparison with field data. We compared our maps with field data at each agro-meteorological site during the period 1985–2020. This involves a total of 3113 on-site verification points, with an average of over 100 verification points per region in 31 provincial-level administrative units in China. The overall accuracy (OA), user accuracy (UA), and producer accuracy (PA) were calculated for each annual maps of rice, wheat, and maize using the monitoring data (Table 2). The OA of the three grain crops was 93.57%. The user accuracy of rice, wheat, and maize were, 92.84%, 96.00%, and 95.36% respectively. Meanwhile, the producer accuracy of rice, wheat, and maize were, 94.98%, 91.45%, and 94.90% respectively.

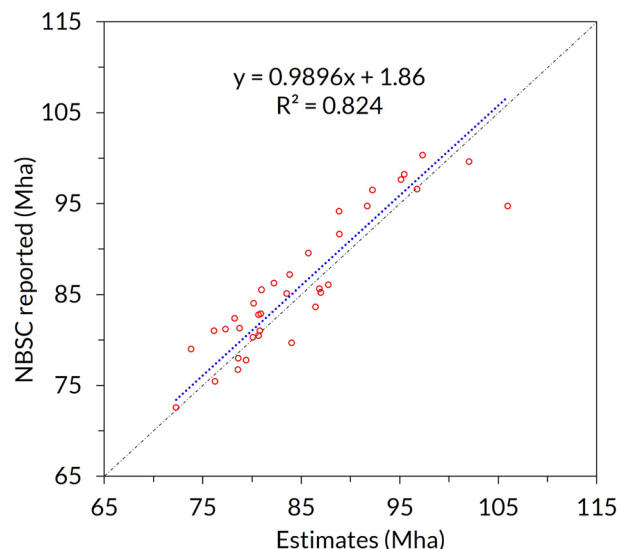


Fig. 6 Time series consistency between MGL total area and statistical data.

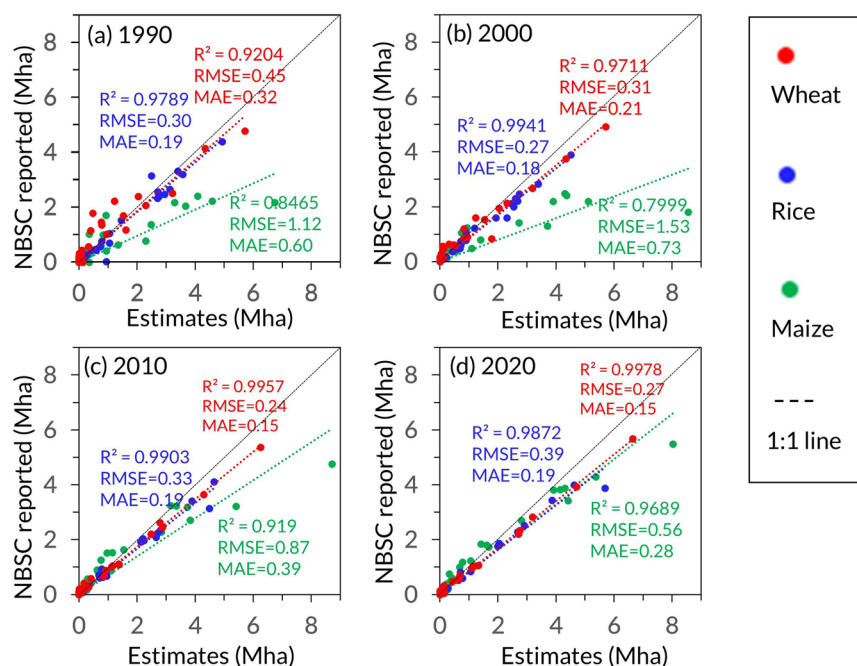


Fig. 7 Spatial validation with statistical data from different provinces.

Validation of MGL total planting area. Here, we have calculated the estimated total MGL planting area in China each year and compared it with the annual major grain planting area reported by NBSC (Fig. 6). The validation results from a long time series show that the MGL dataset has a high consistency with the NBSC reported data ($R^2 = 0.824$). The high consistency also indicates that the mapping accuracy of this dataset is high.

Spatial validation with agricultural statistical data. The plant areas derived from MGL maps were compared with agricultural statistical data reported by NBSC at the provincial level from 1985 to 2020 (Fig. 7). There is a high consistency between the rice planting area of each province and statistical data, with an R^2 close to 1 (0.97–0.99). Meanwhile, the wheat planting areas also agreed well with agricultural statistical data ($R^2 = 0.92$ –1). In contrast, the spatial consistency of maize is lower than that of wheat and rice, but its R^2 is still higher than 0.80 (0.80–0.97). These high levels of consistency also indicate that the MGL dataset has high spatial accuracy.

Comparison of accuracy with existing datasets. To compare the accuracy of the MGL product, we compared the overall accuracy (OA, %) of the mapping results with existing high-resolution crop map using the same sample. The dataset for comparison includes crop maps near 2020 with a resolution of no less than 30

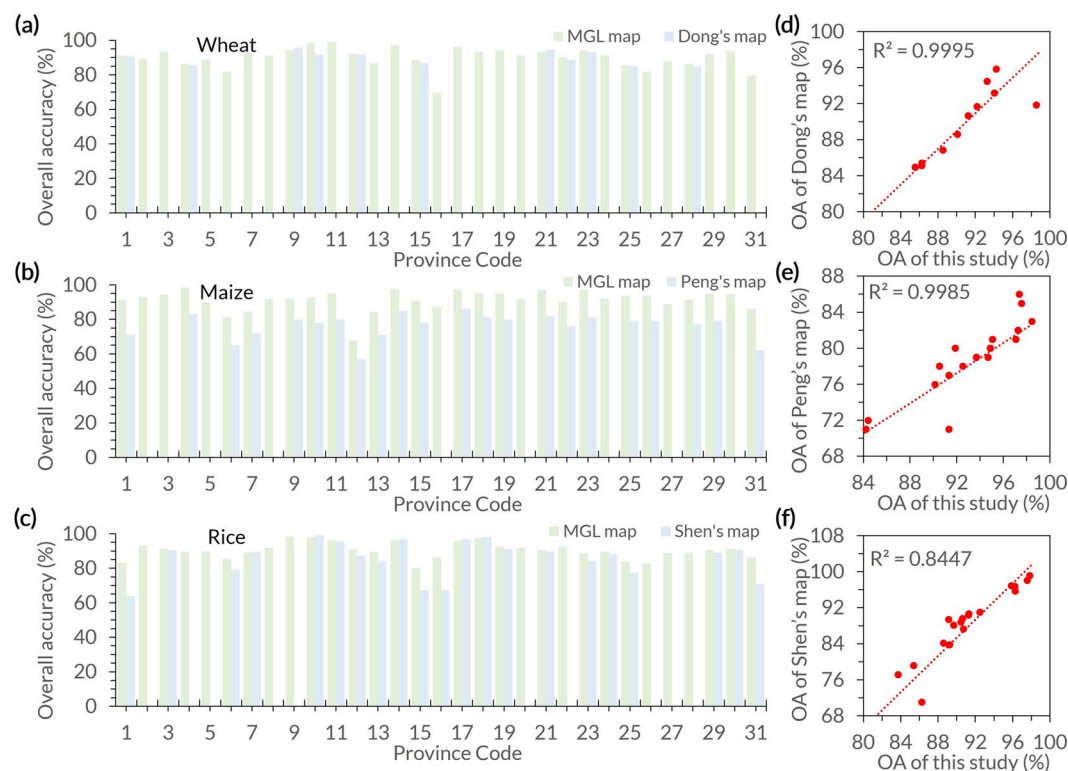


Fig. 8 Comparison of accuracy with existing datasets. (a,b) are the OA comparison and consistency results of wheat. (c,d) are the OA comparison and consistency results of maize. (e,f) are the OA comparison and consistency results of rice. The codes 1–31 on the horizontal axis represent the 31 provincial-level administrative units in China.

meters, including wheat maps in 30 m resolution by Dong *et al.*, maize maps in 30 m resolution by Peng *et al.*, and rice maps in 10–20 m resolution by Shen *et al.* Our wheat map has an OA of 91.34%, slightly higher than that of 89.88% of Dong's wheat map (Fig. 8a). The maize map of this study has an OA of 93.14%, higher than that of 80.06% of Peng's wheat map (Fig. 8b). Similarly, our rice map has an OA of 93.88%, higher than that of 85.23% of Peng's wheat map (Fig. 8c). From a consistency perspective, the MGL results have a high degree of consistency with the mapping OA of a single crop map (Fig. d–f, $R^2 = 0.9995$ in wheat, $R^2 = 0.9986$ in maize, $R^2 = 0.8447$ in rice). These results indicate that the MGL map are slightly higher than similar products with the same resolution available.

Code availability

The code for this study is developed based on GEE. Users can call the link to view the data source code (GEE link: <https://code.earthengine.google.com/d371ab6f6c1bd274c85af270e4ad09c5>). Here is the case code for MGL recognition based on RF algorithm (<https://code.earthengine.google.com/cae24a2d455b2735226c7976b72fc1ba>). The subsequent processing, analysis, mapping, and verification of data are all completed by background computers. These processes are mainly completed in Excel 2021 and ArcGIS 10.2 software, and do not involve code. The full MGL dataset³³ is freely available at <https://doi.org/10.6084/m9.figshare.26212643.v2>.

Received: 22 August 2024; Accepted: 16 December 2024;

Published online: 24 December 2024

References

- Luo, Y. *et al.* Identifying the spatiotemporal changes of annual harvesting areas for three staple crops in China by integrating multi-data sources. *Environ. Res. Lett.* **15**, 074003, <https://doi.org/10.1088/1748-9326/ab80f0> (2020).
- Potapov, P. *et al.* Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century. *Nature Food* **3**, 19–+, <https://doi.org/10.1038/s43016-021-00429-z> (2022).
- Song, X. P. *et al.* Global land change from 1982 to 2016. *Nature* **560**, 639–+, <https://doi.org/10.1038/s41586-018-0411-9> (2018).
- Azzari, G., Jain, M. & Lobell, D. B. Towards fine resolution global maps of crop yields: Testing multiple methods and satellites in three countries. *Remote Sens. Environ.* **202**, 129–141, <https://doi.org/10.1016/j.rse.2017.04.014> (2017).
- Becker-Reshef, I. *et al.* Crop Type Maps for Operational Global Agricultural Monitoring. *Scientific Data* **10**, <https://doi.org/10.1038/s41597-023-02047-9> (2023).
- Yang, Z., Diao, C., Gao, F. & Li, B. EMET: An emergence-based thermal phenological framework for near real-time crop type mapping. *Int. J. Photogramm. Remote Sens.* **215**, 271–291, <https://doi.org/10.1016/j.isprsjprs.2024.07.007> (2024).
- Nabil, M. *et al.* Tree-fruits crop type mapping from Sentinel-1 and Sentinel-2 data integration in Egypt's New Delta project. *Remote Sensing Applications-Society and Environment* **27**, <https://doi.org/10.1016/j.rsase.2022.100776> (2022).

8. Gu, Z. *et al.* Agri-Fuse: A novel spatiotemporal fusion method designed for agricultural scenarios with diverse phenological changes. *Remote Sens. Environ.* **299**, <https://doi.org/10.1016/j.rse.2023.113874> (2023).
9. Liu, X. *et al.* Comparisons between temporal statistical metrics, time series stacks and phenological features derived from NASA Harmonized Landsat Sentinel-2 data for crop type mapping. *Comput. Electron. Agric.* **211**, <https://doi.org/10.1016/j.compag.2023.108015> (2023).
10. Liu, L. *et al.* Detecting crop phenology from vegetation index time-series data by improved shape model fitting in each phenological stage. *Remote Sens. Environ.* **277**, <https://doi.org/10.1016/j.rse.2022.113060> (2022).
11. Hao, P. Y., Zhan, Y. L., Wang, L., Niu, Z. & Shakir, M. Feature Selection of Time Series MODIS Data for Early Crop Classification Using Random Forest: A Case Study in Kansas, USA. *Remote Sensing* **7**, 5347–5369, <https://doi.org/10.3390/rs70505347> (2015).
12. Hu, Q. *et al.* Integrating coarse-resolution images and agricultural statistics to generate sub-pixel crop type maps and reconciled area estimates. *Remote Sens. Environ.* **258**, <https://doi.org/10.1016/j.rse.2021.112365> (2021).
13. Wang, S., Azzari, G. & Lobell, D. B. Crop type mapping without field-level labels: Random forest transfer and unsupervised clustering techniques. *Remote Sens. Environ.* **222**, 303–317, <https://doi.org/10.1016/j.rse.2018.12.026> (2019).
14. Liu, S., Wang, L., Zhang, J. & Ding, S. Opposite effect on soil organic carbon between grain and non-grain crops: Evidence from Main Grain Land, China. *Agric., Ecosyst. Environ.* **379**, 109364, <https://doi.org/10.1016/j.agee.2024.109364> (2025).
15. Press, C. S. Department of Rural Social Economical Survey. *State Bureau of Statistics, China (DRSES-SBS)* (2023).
16. Qiu, B. *et al.* Maps of cropping patterns in China during 2015–2021. *Scientific Data* **9**, 479, <https://doi.org/10.1038/s41597-022-01589-8> (2022).
17. Ray, D. K., Gerber, J. S., MacDonald, G. K. & West, P. C. Climate variation explains a third of global crop yield variability. *Nat. Commun.* **6**, <https://doi.org/10.1038/ncomms6989> (2015).
18. Zhang, G. L. *et al.* Spatiotemporal patterns of paddy rice croplands in China and India from 2000 to 2015. *Sci. Tot. Environ.* **579**, 82–92, <https://doi.org/10.1016/j.scitotenv.2016.10.223> (2017).
19. Shen, R., Peng, Q., Li, X., Chen, X. & Yuan, W. (Science Data Bank, 2024).
20. Pan, B. H. *et al.* High Resolution Distribution Dataset of Double-Season Paddy Rice in China. *Remote Sensing* **13**, <https://doi.org/10.3390/rs13224609> (2021).
21. Dong, J. *et al.* Early-season mapping of winter wheat in China based on Landsat and Sentinel images. *Earth System Science Data* **12**, 3081–3095, <https://doi.org/10.5194/essd-12-3081-2020> (2020).
22. Niu, Q. D. *et al.* A 30 m annual maize phenology dataset from 1985 to 2020 in China. *Earth System Science Data* **14**, 2851–2864, <https://doi.org/10.5194/essd-14-2851-2022> (2022).
23. Macdougall, A. S. *et al.* Widening global variability in grassland biomass since the 1980s. *Nature Ecology & Evolution*, <https://doi.org/10.1038/s41559-024-02500-x> (2024).
24. Maxwell, S. K. & Sylvester, K. M. Identification of “ever-cropped” land (1984–2010) using Landsat annual maximum NDVI image composites: Southwestern Kansas case study. *Remote Sens. Environ.* **121**, 186–195, <https://doi.org/10.1016/j.rse.2012.01.022> (2012).
25. Soriano-González, J., Angelats, E., Martínez-Eixarch, M. & Alcaraz, C. Monitoring rice crop and yield estimation with Sentinel-2 data. *Field. Crop. Res* **281**, <https://doi.org/10.1016/j.fcr.2022.108507> (2022).
26. Hu, J. *et al.* Mapping Diverse Paddy Rice Cropping Patterns in South China Using Harmonized Landsat and Sentinel-2 Data. *Remote Sensing* **15**, <https://doi.org/10.3390/rs15041034> (2023).
27. Jiang, Y. L., Wang, X. H., Huo, M. Y., Chen, F. & He, X. K. Changes of cropping structure lead diversity decline in China during 1985–2015. *J. Environ. Manage.* **346**, <https://doi.org/10.1016/j.jenvman.2023.119051> (2023).
28. Feng, Z. Z. *et al.* Ozone pollution threatens the production of major staple crops in East Asia. *Nature Food* **3**, 47–+, <https://doi.org/10.1038/s43016-021-00422-6> (2022).
29. Qiu, B. W. *et al.* Maps of cropping patterns in China during 2015–2021. *Scientific Data* **9**, <https://doi.org/10.1038/s41597-022-01589-8> (2022).
30. Luo, Y. C. *et al.* Identifying the spatiotemporal changes of annual harvesting areas for three staple crops in China by integrating multi-data sources. *Environ. Res. Lett* **15**, <https://doi.org/10.1088/1748-9326/ab80f0> (2020).
31. Tu, Y. *et al.* A 30 m annual cropland dataset of China from 1986 to 2021. *Earth System Science Data* **16**, 2297–2316, <https://doi.org/10.5194/essd-16-2297-2024> (2024).
32. Xia, X. S. *et al.* Reconstructing Long-Term Forest Cover in China by Fusing National Forest Inventory and 20 Land Use and Land Cover Data Sets. *Journal of Geophysical Research-Biogeosciences* **128**, <https://doi.org/10.1029/2022jg007101> (2023).
33. Liu, S., Wang, L. & Zhang, J. The dataset of main grain land changes in China over 1985–2020. *Figshare* <https://doi.org/10.6084/m9.figshare.26212643.v2> (2024).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No:42401082), Postdoctoral Fellowship Program of CPSF (Grant No: GZC20232757; GZC20240855), Project funded by China Postdoctoral Science Foundation (Grant No: 2024M763351; 2024M751623), and the Special Research Assistant Program of the Chinese Academy of Sciences. Also, authors gratefully acknowledge the colleagues and friends in Aerospace Information Research Institute for offering the basic data and relevant information. We also want to thank the anonymous reviewers for their valuable suggestions and comments that contributed to improving the quality of the manuscript.

Author contributions

Shidong Liu designed the study and the methodology. Shidong Liu and Li Wang wrote the code and implemented the algorithms. Shidong Liu and Jie Zhang composed the draft of this manuscript and analyzed the data and edited the manuscript.

Competing interests

The authors declare that they have no conflicts of interest with this work. We declare that we do not have any commercial or associated interest that would represent a conflict of interest in connection with the work submitted.

Additional information

Correspondence and requests for materials should be addressed to L.W. or J.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024