Journal of Hydrology 613 (2022) 128413

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Integrating remote sensing, irrigation suitability and statistical data for irrigated cropland mapping over mainland China

Ling Zhang [a,b,*], Kun Zhang [c,d], Xiufang Zhu [e], Hao Chen [a,b], Weizhen Wang [a]

[a] Key Laboratory of Remote Sensing of Gansu Province, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China
[b] Key Laboratory of Land Surface Process and Climate Change in Cold and Arid Regions, Chinese Academy of Sciences, Lanzhou 730000, China
[c] School of Biological Sciences, The University of Hong Kong, Hong Kong, China
[d] Department of Mathematics, The University of Hong Kong, Hong Kong, China
[e] State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and Institute of Remote Sensing and Digital Earth of Chinese Academy of Sciences, Beijing 100875, China

ARTICLE INFO

This manuscript was handled by Jiri Simunek, Editor-in-Chief, with the assistance of Zailin Huo, Associate Editor

Keywords:
Irrigation
Irrigated cropland
Remote sensing
Irrigation suitability
Machine learning
China

ABSTRACT

Knowledge of irrigation location and extent is essential for irrigation-water use estimation and water resource management. However, it remains a great challenge to map irrigated areas at large spatial scales due to the great variation in climate, geography, and agricultural practices, as well as the lack of sufficient ground truth data. This study proposed a novel approach to develop the first 250-m irrigated cropland map in mainland China (CIrrMap250) by integrating remote sensing, irrigation suitability, and irrigated area statistics. We assessed the performance of CIrrMap250 and compared it with three irrigation maps (i.e., EVI-map, NDVI-map, GI-map) generated using the threshold-based classification method and four other existing maps, including GMIA2005 (Siebert et al., 2005), GIAM2000 (Thenkabail et al., 2009), Zhu-map (Zhu et al., 2014), and Meier-map (Meier et al., 2018). Results indicate that CIrrMap250 and all other maps capture well the intensively irrigated areas such as the North China Plain and Northwest China, as well as many large-scale irrigation districts. However, all maps except CIrrMap250 tend to underestimate irrigated cropland in river valleys while overestimating irrigated cropland in the mountainous areas, as illustrated by the field-surveyed irrigation districts, due to the neglect of the mixed grid effects. Compared to other irrigation maps, CIrrMap250 exhibits a better agreement with the reference points, achieving improvements in Kappa coefficient and overall accuracy by 8% up to about 2 times. The irrigated area estimates of CIrrMap250 are very close to the statistical data due to their usage in generating the training pool. Further analysis indicates CIrrMap250 has a greater proportion of irrigated cropland at lower elevations, on smaller slopes, and near water bodies than the other maps. There is large uncertainty in irrigation ratio estimates due to the varying cropland area from multiple sources. This study demonstrates the effectiveness of the new irrigation mapping method and highlights the great potential of combining irrigation suitability with remote sensing and statistical data to improve the accuracy of large-scale irrigated cropland mapping.

1. Introduction

Irrigation greatly enhances agricultural yields and plays a critical role in safeguarding food security (Wang et al., 2021). Irrigated agriculture contributes to ~ 40 % of global food production in just 20 % of cropland, but accounts for ~ 70 % global freshwater withdrawals and 90 % of consumptive water use (Siebert and Döll, 2010; Wada et al., 2013). Irrigation is expected to expand in the future due to climate change, growing food demand, and agricultural intensification (Deines et al., 2017). Extensive irrigation has caused dramatic changes in water cycles, putting unprecedented pressure on sustainable freshwater use (McDermid et al., 2021). Reconciling irrigation benefits with water stress and environmental impacts remains a grand challenge for managers and policy makers in the twenty-first century (Salmon et al., 2015; Rosa et al., 2020).

Irrigation-water withdrawal is mostly driven by irrigated area (Puy et al., 2021). Knowledge of irrigation distribution is a first step towards irrigation water estimation and water resource management (Portmann

* Corresponding author at: Key Laboratory of Remote Sensing of Gansu Province, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China.
E-mail address: zhanglingky@lzb.ac.cn (L. Zhang).

https://doi.org/10.1016/j.jhydrol.2022.128413
Received 22 November 2021; Received in revised form 29 July 2022; Accepted 21 August 2022
Available online 28 August 2022

et al., 2010; Xie and Lark, 2021). However, irrigation location and extent are poorly understood, especially at national to global scales, due to the hidden nature of irrigation signals and the limited ground truth data available for training classification algorithms (Ozdogan and Gutman, 2008; Liu et al., 2018; Xie et al., 2021). Insufficient information on irrigation distribution is unfavorable for sustainable water use and management (Nagaraj et al., 2021) and will introduce large uncertainties and biases in the modeling of surface and hydrologic processes, land–atmosphere interactions, and crop growth.

Satellite remote sensing provides unprecedented opportunities to detect and map irrigated areas at multiple spatial scales. Various methods have been proposed to identify irrigated areas using remote sensing data, which can be classified into two categories, namely optical and near-infrared methods and microwave-based methods (Massari et al., 2021). Visual interpretation is perhaps the most primitive method that use optical and near-infrared signatures to detect irrigated area, but is a cost and time intensive procedure (Karthikeyan et al., 2020). Besides the spectral signatures, many vegetation indices such as the normalized difference vegetation index (NDVI) (Rouse et al., 1974), normalized difference water index (NDWI) (Gao, 1996), and green index (GI) (Gitelson, 2005) have been used to map irrigated areas, the rationale behind which is that irrigation reduces crop water stresses and increase productivity, resulting in higher greenness and water content of irrigated crops than rainfed crops (Nagaraj et al., 2021). These vegetation indices are often used with classification algorithms trained with ground-truth data, such as decision trees (Ozdogan and Gutman, 2008; Ambika et al., 2016b), support vector machine (Sharma et al., 2021), and random forest (Peña-Arancibia et al., 2014; Deines et al., 2017), to distinguish between irrigated and rainfed croplands. In some exceptions, the vegetation indices are compared to the target spectra of ground samples (Lu et al., 2021), adjacent forest pixels (Xiang et al., 2019), and time-series precipitation to detect irrigation signals (Chen et al., 2018). In recent years, the microwave-based irrigation mapping with satellite soil moisture products has also attracted attention (Dari et al., 2021). The rationale behind the microwave-based approach is that irrigation increases soil moisture and causes changes in soil moisture and radar backscatter signals. Thus, microwave information is usually compared with simulations under "natural" conditions to detected irrigated areas (Kumar et al., 2015; Zaussinger et al., 2019; Zohaib et al., 2019) or used with the classification algorithms trained with ground truth data (Gao et al., 2018; Pageot et al., 2020; Bazzi et al., 2021).

Remote sensing-based irrigated cropland extraction methods have shown high accuracy and reliability in small-scale areas (e.g., watersheds and irrigation districts). Nevertheless, it is still a great challenge to map irrigated area at large spatial scales (e.g., national and global scales) due to the wide variation in climate, geography and agricultural practices (Salmon et al., 2015), as well as the lack of adequate ground truth data (Ozdogan et al., 2010; Xie and Lark, 2021). There have been several attempts to map the global irrigated area. Siebert et al. (2005) generated the global map of irrigated areas (GMIA) by combining national and subnational statistics with geospatial information on irrigation location and extent. Thenkabail et al. (2009) derived the global irrigated area map (GIAM) from remote sensing using the spectral matching technique and decision tree algorithm. Later, Salmon et al. (2015) fused remote sensing classification results with agricultural inventory data to generate a map of global rainfed, irrigated, and paddy croplands (GRIPC). Meier et al. (2018) merged the downscaled GMIA map with the remote sensing-derived irrigation map to generate a new global high-resolution irrigation map. In addition, several national irrigation maps have been produced in recent years, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) Irrigated Agriculture Dataset (MIrAD-US) (Pervez and Brown, 2010), Landsat-based Irrigation Dataset across the conterminous United States (LANID-US) (Xie et al., 2019; Xie and Lark, 2021), and remotely sensed high resolution irrigated area map in India (Ambika et al., 2016b). Typically, national irrigation maps are more accurate than global maps due to the availability of more ground data, higher spatial resolution, and reliable auxiliary information. China is a large agricultural country with the largest irrigated area (65.87 million ha) in the world, followed by India (62.00 million ha) and the United States (26.71 million ha) (IDCD, 2018). However, to our knowledge, few studies have mapped irrigated areas throughout China (Xiang et al., 2020; Zhu et al., 2021; Zhang et al., 2022), and the spatial resolution of irrigation map is far below that of other countries (e.g. the United States and India). This limits the accuracy of irrigation-water use estimates in China and significantly affects irrigation-related ecological, hydrological, and climatic studies.

To address the above gaps, this study proposes a novel approach to develop the first 250-m irrigated cropland map in mainland China (CIrrMap250) for the year 2000 by integrating remote sensing, irrigation suitability, and irrigated area statistics. The major difference between our approach and other existing methods lies in the incorporation irrigation suitability into the mapping process. Our objectives of are threefold: (1) to test the effectiveness of the new irrigation mapping method; (ii) to compare the performance of CIrrMap250 with three irrigation maps generated using the traditional threshold-based method and four other existing maps (Siebert et al., 2005; Thenkabail et al., 2009; Zhu et al., 2014; Meier et al., 2018); and (iii) to reveal the spatial distribution features of irrigated cropland in mainland China.

2. Study area

China lies between 15°-50°N and 65°-135°E and covers an area of approximately 9.6 million square kilometers. This study was carried out in mainland China (Fig. 1), which is characterized by a wide range of elevations, from 176 m below sea level at Ayding Lake in the Turpan depression of Xinjiang to 8,794 m above sea level on the Qinghai Tibetan Plateau. According to the Köppen-Geiger climate classification (Biggs et al., 2006), southeastern China has a warm temperate climate; northwestern China has an arid climate; northeastern and parts of central China have a cold climate; and the high-altitude Qinghai Tibetan Plateau has a polar and cold climate. Cropland is extensively distributed across mainland China, with a greater concentration in the eastern, northeastern, and central regions than in the arid northwestern region, particularly the North China Plain, Northeast China Plain, Guanzhong Plain, Sichuan Basin, and Middle and Lower Yangtze River Plain. As shown in Fig. 1**b**, the study area is divided into 3,652 mapping units. There are 686 reference points covering mainland China, which are used to evaluate the accuracy of the irrigation maps.

3. Materials and methods

The workflow of the study is summarized in Fig. 2. We first collected and processed three types of data, including satellite data, irrigated area statistics, and auxiliary data. Then, we combined the satellite-derived vegetation indices (NDVI, EVI and GI) with irrigated area statistics from 3,625 administrative units to drive three irrigation maps (i.e., NDVI-map, EVI-map, and GI-map) using the threshold-based classification method (Pervez and Brown, 2010). Afterward, we produced another irrigated cropland map (i.e., CIrrMap250) using the newly proposed semiautomatic classification approach. The new approach consists of three major steps, including irrigation suitability analysis, training pool generation based on suitability-adjusted vegetation indices and irrigated area statistics, and machine learning-based classification of irrigated and rainfed cropland. Lastly, we evaluated the accuracy of CIrrMap250, NDVI-map, EVI-map, GI-map, and four other existing maps (i.e., GMIA2005, GIAM2000 Zhu-map, and Meier-map) from three different perspectives, including the distribution of irrigated croplands, agreement with reference points, and consistency with irrigated area statistics.

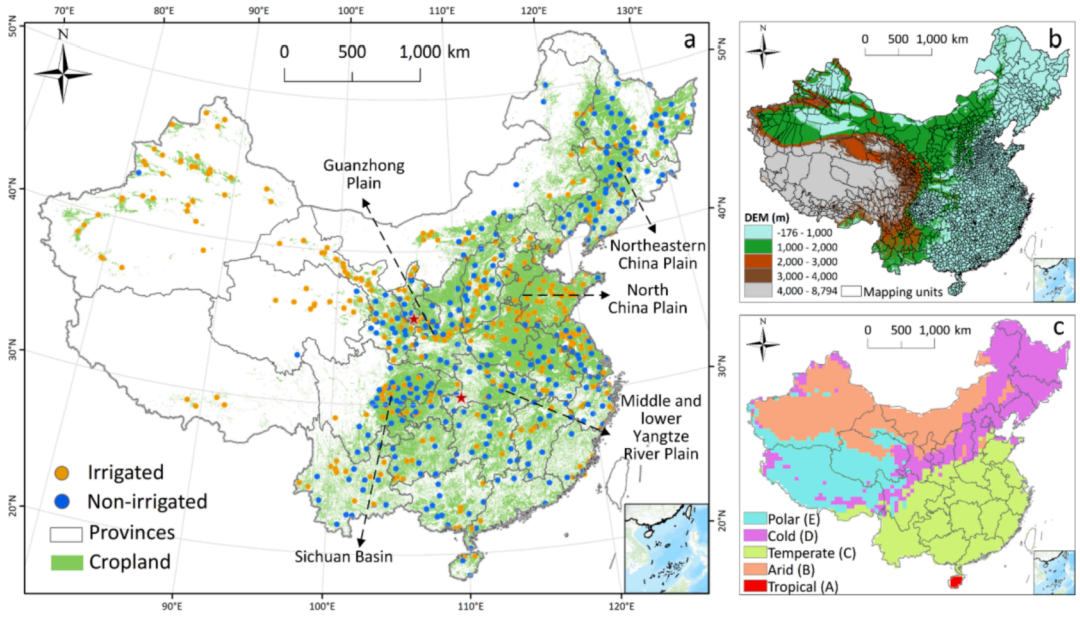

Fig. 1. Overview of the study area. a, Spatial distribution of croplands and validation points used for accuracy assessment. b, Topography and mapping units across mainland China. Digital Elevation Model (DEM) with a spatial resolution of 90 m was obtained from the Shuttle Radar Topography Mission (SRTM). c, Köppen-Geiger climate classification in mainland China.

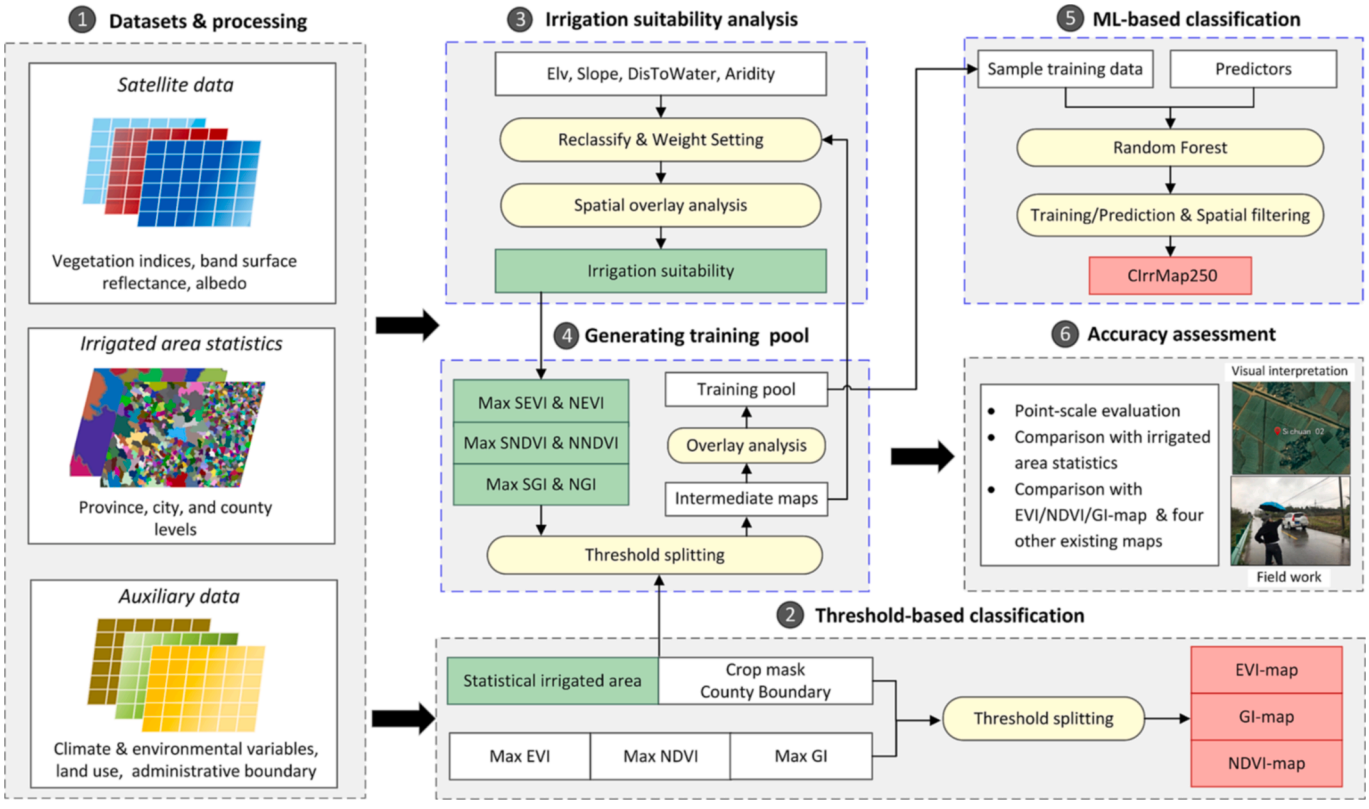

Fig. 2. Workflow of the study. The number and gray boxes indicate the major steps involved in this study. White boxes are the input data sets or variables; the green boxes are the three major data sources used for irrigated cropland mapping; the yellow boxes are the analysis methods or tools and the red boxes are the irrigation maps generated in the study. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.1. Datasets and processing

3.1.1. Satellite data

The MODIS vegetation indices (MOD13Q1) Version 6 data including two vegetation layers, i.e., NDVI and enhanced vegetation index (EVI) (Huete et al., 1997), were used in the study. As shown in Table 1, NDVI and EVI are 16-day composite products with a spatial resolution of 250 m. Meanwhile, the 500-m and 8-day surface spectral reflectance of band 04 provided by MOD09A1 was resampled to 250 m through nearest neighbor interpolation (Debeurs and Townsend, 2008). It was used to

Table 1
Summary of the MODIS-derived vegetation indices used in this study.

Vegetation indices	Formula	MODIS bands	Resolution
NDVI	(NIR - Red) / (NIR + Red)	Bands 01, 02	250 m/16 day
EVI	2.5*(NIR-Red) / (NIR + 6*Red–7.5*Blue + 1)	Bands 01, 02, 03	250 m/16 day
GI	NIR/Green	Bands 01, 04	250 m/8day
NNDVI	Max annual NDVI / median neighborhood value	–	250 m/ annually
NEVI, NGI	Similar to the calculation method of NNDVI	–	250 m/ annually

Red: band 01, Blue: band 03, near-infrared (NIR): band 02, Green: band 04.

derive another vegetation index (i.e., GI) along with the 250-m and 8-day surface reflectance of band 01 provided by MOD09Q1. All MODIS data were quality filtered using quality and usefulness indicators, and only those cloud and snow/ice free pixels with decreasing to highest quality were considered to be reliable (Hilker et al., 2012). Unreliable pixels were reconstructed using a simple nearest neighbor interpolation method. We calculated the maximum, minimum, range, median values of vegetation indices for each crop grid. Note that the maximum and minimum values were calculated at the 95th and 10th percentile, respectively, to exclude the unidentified cloudy and poor-quality values (Xie and Lark, 2021). To better account for crop and climate gradients, we further derived a set of neighborhood normalized vegetation indices (NNDVI, NEVI, NGI) (Deines et al., 2019) by dividing the annual peak vegetation index of a target pixel to the median value of its neighborhood pixels within a 25 km radius circular kernel.

3.1.2. Irrigated area statistics

The irrigated area statistics were compiled through a top-down and bottom-up approach. The county-level and city-level irrigated area were first collected from the Provincial Statistical Yearbook, the Rural Statistical Yearbook, and the China Statistical Yearbook for Regional Economy (National Bureau of Statistics of China, 2001b). In China, municipal administration (city) consists of districts and counties, and here we referred to them all as counties. Counties without any irrigated area were merged into their nearest neighbors. We aggregated the irrigated areas at the county level to the city level and compared them with the city-level statistics for consistency checks. Where inconsistencies existed, we further collected irrigation-water use data from the provincial water resource bulletin, and irrigation areas with a high correlation to irrigation-water use were considered reliable for subsequent analysis. The city-level irrigated areas were finally aggregated to the provincial level and compared them with the values reported by the China Water Statistical Yearbook (CWSY) (China Ministry of Water Resources, 2001). We put higher confidence in the CWSY data because it is published by the Chinese Ministry of Water Resources and edited by water practitioners and researchers throughout China. As shown in Supplementary **Fig. S1**, the city-aggregated irrigated areas are very close to the CWSY statistics with a coefficient of determination (R^2) of 0.99. For provinces where inconsistencies exist, we adjusted the county-level irrigated area using Eq. (1).

$$IrrArea_{i,j} = \frac{IrrArea_{CWSY,j}}{IrrArea_{agg,j}} \times IrrOrgin_{i,j} \quad (1)$$

where $IrrArea_{i,j}$ and $IrrOrgin_{i,j}$ are the adjusted irrigated area and the original irrigated area in county i of province j, respectively; $IrrArea_{CWSY,j}$ and $IrrArea_{agg,j}$ are the CWSY-reported irrigated area and the city-aggregated irrigated area of province j, respectively. County-level irrigated area statistics are available for>80 % of provinces in mainland China (Supplementary **Fig. S2**), and we totally have 3,652 mapping units with irrigated area statistics (Fig. 1**b**). Note that Xinjiang, Heilongjiang and Fujian provinces have some irrigated croplands managed by the Production and Construction Corps, which we merged into the cities where they are located by population size following the method of Zhu et al. (2014).

3.1.3. Auxiliary data

Various auxiliary data were used in this study, including climatic and environmental variables, land use and land cover, and administrative boundaries. Climate data, including precipitation, temperature, pressure, wind speed and downward solar radiation, were extracted from the China Meteorological Forcing Dataset (He and Yang, 2016; He et al., 2020). These data were used in conjunction with the MCD43A3 albedo product to estimate potential evapotranspiration (PET) using the Priestley-Taylor method (Priestley and Taylor, 1972), as well as the aridity index defined as the ratio of precipitation to PET. Environmental data include elevation, slope, soil type, and distance to water bodies. The elevation data was taken from the Shuttle Radar Topography Mission digital elevation model (i.e., SRTM DEM); and the slope map was generated from the SRTM DEM data using the slope function in ArcGIS software. Distance to water bodies was calculated from the spatial distribution data of water bodies, including rivers, lakes, reservoirs, canals, and ponds, using the Euclidean distance tool in ArcGIS software. The 2000 land use/cover (LULC) map with a spatial resolution of 30 m was sourced from the Land Use Status Remote Sensing Monitoring Database of China (Xu et al., 2018). The map includes six major LULC types, namely, cropland, forest, grassland, water body, urban land, and unused land. The spatial distribution and proportion of each land use/cover type at 250 m resolution was estimated from the original 30-m LULC map. All data used in this study were obtained partly from the National Tibetan Plateau (https://data.tpdc.ac.cn/zh-hans/) and partly from the Resource and Environment Science and Data Center (https://www.resdc.cn/Default.aspx).

3.2. Threshold-based classification method

Threshold-based classification is a simple yet robust, easy-to-implement approach for creating irrigation maps that assumes the existence of thresholds for certain indices that can differentiate between irrigated and rainfed cropland. These indices can be the annual peak vegetation index (Pervez and Brown, 2010), the irrigation potential index (Zhu et al., 2014), or the LSWI difference index (Xiang et al., 2020); and the thresholds can be calibrated based on irrigated area statistics, ground truth data, or experience from previous studies (Pervez and Brown, 2010; Shahriar Pervez et al., 2014; Meier et al., 2018). Here, following Pervez and Brown (2010), we classified irrigated and rainfed cropland using the thresholds of annual peak vegetation indices (NDVI, EVI, and GI). Specifically, for each mapping unit of mainland China, we first sorted the annual peak vegetation indices of all croplands in descending order and estimated the cumulative irrigated area sequentially. The accumulated area was then compared with the statistical irrigated area; and the value of vegetation index for the grid where the cumulative irrigated area was close to the statistical data was determined as the threshold. We lastly classify the cropland into "irrigated" and "rainfed" by comparing annual peak vegetation index with the determined threshold, as in Eq. (2).

$$cropland_{i,j} = \begin{cases} irrigated_{i,j} & PVI \geq threshold \\ rainfed_{i,j} & PVI < threshold \end{cases} \quad (2)$$

where PVI are the annual peak vegetation index of the crop grid i in the mapping unit j over mainland China. Classification results using NDVI, EVI and GI were referred to as NDVI-map, EVI-map, and GI-map, respectively, hereafter. The annual minimum, median, and range of NDVI, EVI, and GVI were also tested, but showed obviously lower performance than the annual peak values in mapping irrigated cropland.

3.3. Semiautomatic classification approach

3.3.1. Irrigation suitability analysis

Irrigation suitability analysis has been widely used to assess potential land suitable for irrigation (Chen et al., 2010; Worqlul et al., 2015; Worqlul et al., 2017; Li and Chen, 2020), but has not received much attention in the irrigation mapping. In this study, we used the multi-indicator evaluation (MCE) technique to assess the irrigation suitability of croplands in mainland China. Our study considered elevation, slope, distance to water bodies (DistToWater), and drought index, as these natural and geographic conditions have proven to be important factors controlling the spatial distribution of irrigated cropland in China (Liu et al., 2022). Based on our experience and field work, as well as the findings of previous studies, we hypothesize that croplands with smaller DistToWater, lower elevation and slope, and higher aridity index have greater irrigation suitability and potential. It is easy to understand that the closer to water bodies, the greater the possibility of irrigation, since closer proximity to water bodies means easier access to irrigation water. Some previous studies have even selected the irrigation reference points by choosing croplands close to water bodies (Dong et al., 2009; Zhang et al., 2022). The higher the elevation, the more difficult it is to access water resources, leading to a lower likelihood of irrigation (Cuo et al., 2013). For example, in our field work, we found that in the high-elevation areas of the Loess Plateau, residents have access to domestic water only through deep wells, while crop growth is completely dependent on rainfall. Meanwhile, agricultural productivity is low at high altitudes in China due to lack of irrigation and higher transportation and labor costs (Li et al., 2015). In addition, it was found that paddy fields, which are usually equipped for irrigation, tend to be clustered in the lowlands of China compared to drylands (Liu et al., 2005). Areas with greater slopes have poor water holding capacity and are unfavorable for irrigation facilities (Akıncı et al., 2013; Worqlul et al., 2017; Mandal et al., 2018). Generally, areas with slopes exceeding 8 % are considered infeasible for any surface irrigation system (Shahriar Pervez et al., 2014; Ambika et al., 2016a). Therefore, it is widely believed that the gentler the slope, the more likely is the existence of irrigated cropland (Yin et al., 2020; Ishikawa and Yamazaki, 2021). Croplands with higher aridity indices (i.e., less precipitation but higher PET) are also more likely to be irrigated due to the greater demand for irrigation water (Xu et al., 2018; Yin et al., 2020).

Elevation, slope, DistToWater, and aridity index were first reclassified into different suitability types, as shown in Supplementary **Table S1**. The elevations of the crop grids were reclassified into four types using the dividing points of 100, 300, 500 m above the lowest elevation of the mapping unit. There are also four suitability types for both Slope and DistToWater. The dividing points for slopes are 2 %, 4 %, and 8 %, while dividing points for DistToWater are 1,000, 10,000, and 20,000 m. Using the dividing points, aridity indices were reclassified into ten suitability types. Afterwards, the above reclassified factors were assigned with different suitability values, with higher values given to areas with lower elevation, smaller slope and DistToWater, and higher aridity index. Lastly, we estimated the overall irrigation suitability of the croplands by combining the suitability values of elevation, slope, DistToWater, and aridity index, as in Eq. (3).

$$S_{i,j,k} = \frac{1}{4}w_{1,k}SElev_{i,j} + \frac{1}{4}w_{2,k}SSlope_{i,j} + \frac{1}{4}w_{3,k}SDW_{i,j} + \frac{1}{10}w_{4,k}SArid_{i,j} \quad (3)$$

where $S_{i,j,k}$ is the irrigation suitability for cropland i in mapping unit j of province k; w is the weight of the influencing factors; *SElev*, *SSlope*, *SDW*, and *SArid* are the suitability values of elevation, slope, DistToWater, and aridity index, respectively. The weight setting is critical for irrigation suability analysis. In this study, weights were set in each province using the trial-and-error method based on the irrigated cropland distribution in the intermediate maps (See next section). According to our field work and experience, appropriate weights should be given so that distribution of irrigated cropland have the following key characteristics: (i) irrigated croplands are mostly distributed near water bodies, at lower elevations, and on smaller slopes; (ii) irrigated croplands in plain areas are relatively evenly distributed; (iii) irrigated croplands are generally continuously distributed due to the influence of water facilities; and (iv) irrigated croplands in non-plain regions are mainly distributed in river valleys.

3.3.2. Generating training pool

Inspired by the work of Xie et al. (2019), we designed a threshold-calibrated method to automatically generate the training pool (i.e., potential training data) for each mapping unit of mainland China. We first proposed a new irrigation suitability-adjusted vegetation index (SVI)with the assumption that irrigated cropland is not only greener, but also more suitable for irrigation than rainfed cropland. The SVI is defined in Eq. (4).

$$SVI = S \times PVI \quad (4)$$

where S is the irrigation suability; and *PVI* is the annual peak vegetation index. The suitability-adjusted NDVI, EVI and GI were named as SNDVI, SEVI and SGI. Three intermediate irrigation maps were then generated based on SNDVI, SEVI, and SGI using the threshold-based classification method (Section 3.2). Note that the annual peak NNDVI, NEVI, and NGI were used in North China plain, considering that crop gradient is not negligible, but the gradients in elevation, slope and aridity index are small and groundwater is widely used for irrigation. Furthermore, cropland is discontinuously distributed in many parts of mainland China, and many croplands are mixed with other land use/cover types within the 250-m grids. The *PVI* of the cropland-urban/water mixed grids would be underestimated compared to the pure cropland pixels, while the cropland-forest/grassland mixed grids would be overestimated. We recalculated the *PVI* for the mixed pixels with the cropland proportion less than 90 % from the nearest pure cropland pixel. In the case of few pure cropland pixels over mountainous areas, the *PVI* values of mixed pixels were adjusted by the proportion of croplands, considering that irrigation is more likely to occur on continuously distributed croplands. Lastly, the intermediate irrigation maps were overlaid, and pixels that were simultaneously identified as irrigated or rainfed cropland by these maps were used as the training pool for further classification.

3.3.3. Classification using random forest

We used the Random Forest (RF) (Breiman, 2001) to classify irrigated and rain-fed cropland using the random samples from the training pool. For each mapping unit of mainland China, we randomly sample 200 rainfed cropland grids and 200 irrigated cropland grids following Xie et al. (2019) to balance the need for adequate samples and computational efficiency. The sample size was increased to 4,000 (i.e., 2,000 each for irrigated and rainfed cropland) for the mapping units only with city-level irrigated area statistics. Sensitive analyses indicated the use of different sets of training samples has minimal effect on the classification accuracy (Supplementary **Fig. S3**). We implemented the RF algorithm using the MATLAB TreeBagger function, and the hyperparameters (Table 2) were determined through a trial-and-error procedure (Zhang et al., 2021). The selected predictors include vegetation indices (i.e., SEVI, NEVI, SNDVI, NNDVI, SGI and NGI), climatic variables (i.e., precipitation, temperature, wind speed, solar radiation, and aridity index), and environmental variables (i.e., latitude, longitude, crop intensity,

Table 2
Calibrated hyperparameters of the RF algorithm.

Hyperparameters	Descriptions	values
Ntree	Number of trees	100
MinObs	Minimum number of observations per node	10
Nsplit	Number of variables randomly sampled at each decision split	7

elevation, DistToWater, slope, and soil type). The RF-based classification of irrigated and rainfed cropland was carried out separately in each mapping unit. After classification, we applied a spatial filter (i.e., 7 × 7 window) to remove isolated pixels (less than5% of the window area) and detect the missed irrigated croplands (>95 % of the window area), resulting in the final irrigated cropland map in mainland China with a spatial resolution of 250 m (i.e., CIrrMap250). We also tested two additional commonly used machine learning algorithms (i.e., artificial neural network (Gardner and Dorling, 1998) and extreme learning machine (Huang et al., 2012)) for classification of irrigated and rainfed croplands, but they showed lower accuracy than RF (Supplementary **Fig. S4**).

3.4. Accuracy assessment

The accuracy of CIrrMap250 were assessed from three different aspects. First, the spatial distribution of irrigated cropland in CIrrMap250 was compared with EVI-map, NDVI-map, and GI-map and four other existing maps including GMIA 2005 (Siebert et al., 2005), GIAM2000 (Thenkabail et al., 2009), Zhu-map (Zhu et al., 2014), and Meier-map (Meier et al., 2018). The performance of these maps was evaluated by analyzing their ability in detecting intensively irrigated areas and large-scale irrigation districts of mainland China, as well as two field-surveyed irrigation districts. Second, we further conducted a point-scale evaluation of CIrrMap250 and other irrigation maps using 686 reference points (Fig. 1). Most of the reference points (~90 %) were sourced from Zhu et al. (2014), which were partly collected from the crop growth and soil moisture dataset provided by the China Meteorological Data Sharing Service System (https://data.cma.cn/) and partly obtained using Google Earth and irrigation information on large irrigated districts provided by the China Irrigation and Drainage Development Center (https://www.jsgg.com.cn/ciddc/Index.asp). The remaining 10 % of reference points were derived through our field work and visual interpretation of Google satellite maps. As listed in Table 3, the performance metrics including Kappa coefficient, overall accuracy, and producer's accuracy were used for quantitative assessments. Finally, we compared the irrigated area statistics with the estimates from CIrrMap250, Zhu-map, GMIA2005, and Meier-map at different spatial scales.

4. Results and discussion

4.1. Spatial distribution of irrigated cropland in CIrrMap250

Fig. 3 depicts the spatial distribution of irrigated and rainfed cropland in CIrrMap250. We can see that irrigated croplands are mainly distributed in the North China Plain and Northwest China. Meanwhile, many large-scale irrigation districts such as Hetao, Qingtongxia, Fenhe, Baojixia, Dujiangyan, Zhanghe, and Pishihang also have intensively irrigated croplands. At the provincial level, Shandong holds the largest irrigated area (4,940 Kha), followed by Henan (4,834 Kha), Hebei (4,566 Kha), Jiangsu (3,999 Kha), Anhui (3,236 Kha), and Xinjiang (3,172 Kha). The provinces including Tibet, Qinghai, Tianjin, Beijing, Shanghai, and Hainan have relatively smaller areas of irrigated croplands than other provinces in mainland China. At the basin scale, the Yangtze River basin has the largest irrigated area (14,939 Kha), followed by the Huai River basin (10,622 Kha), Hai River basin (7,400 Kha). The Southwest River basin has the lowest irrigated cropland area among the nine river basins in mainland China.

Table 3
Definitions of the performance metrics.

Metrics	Formula	Variables
Overall accuracy	$\frac{\sum_{i=1}^{n} P_{ii}}{N}$	n is the number of classes; P_{ii} is the number of pixels on row i and column i in the confusion matrix, which represent the total number of pixels correctly classified; N is total number of pixels used for accuracy evaluation; P_{i+} and P_{+i} are the total number of pixels on row i and column i, respectively.
Kappa coefficient	$\frac{N\sum_{i=1}^{n} P_{ii} - \sum_{i=1}^{n} P_{i+} \times P_{+i}}{N^2 - \sum_{i=1}^{n} P_{i+} \times P_{+i}}$	
Producer's accuracy	$\frac{P_{ii}}{P_{+i}}$	

4.2. Comparison of irrigated cropland distribution

Fig. 4 compares the distribution of irrigated croplands in different irrigation maps. Intensively irrigated areas such as the North China Plain, Northwest China, and large-scale irrigation districts are overall well captured by all the maps. The spatial pattern of irrigated croplands in EVI-map, NDVI-map, GI-map is similar to CIrrMap250 due to the consistent constraints of irrigated area statistics on the mapping results. However, a closer look reveals that the distribution of irrigated cropland in the EVI/NDVI/GI-map is more uniform and extensive in the southern and central regions of mainland China than in CIrrMap250. There are notable differences between CIrrMap250 and the four existing maps at both regional and local scales. Compared with other irrigation maps, irrigated cropland in Zhu-map is less distributed in central and southeastern Chongqing, northeastern Heilongjiang, and the Baojixia irrigation district, but more distributed in northern Hebei province. GIAM2000 has more irrigated areas in the North China Plain, Northeast China Plain, and Sichuan and Guangxi provinces, but obviously less irrigated area in provinces such as Tibet, Qinghai, Gansu, and Guangdong. GAIM2005 does not provide much details on the distribution of irrigated cropland due to the coarse spatial resolution. Irrigated cropland in GAIM2005 appears to be more distributed in northeastern China, southern Sichuan Province, and northern Shaanxi than the other irrigation maps, but less distributed in Jiangxi, Jiangsu, Sichuan, and Hubei.

The two panels inserted in each subplot of Fig. 4 show the distribution of irrigated cropland in two field-surveyed irrigation districts, i.e., the Jinghe Irrigation District (upper panel) and the Xuhuiqu Irrigation District (lower panel). Both Jinghe and Xuhuiqu irrigation districts are located in river valleys where irrigated cropland is concentrated. CIrrMap250 captures well the concentration of irrigated cropland in these two irrigation districts. However, all other maps tend to underestimate irrigated cropland in the Jinghe and Xuhuiqu irrigation districts while overestimating irrigated cropland in the surrounding mountainous areas. The similar results can be extensively observed if we enlarge EVI-map, NDVI-map, GI-map, Zhu-map, GIAM2000, GMIA2005, and Meier-map.

4.3. Point-scale accuracy evaluation

Fig. 5 summarizes the point-scale accuracy of the different irrigation maps. GAIM2005 is not evaluated here because of its coarse spatial resolution. We can see that CIrrMap250 attains a Kappa coefficient of 0.58 in mainland China, obviously higher than EVI-map (0.47), NDVI-map (0.44) and GI-map (0.46), as well as the other three existing maps (0.20 ~ 0.28). The improvement in Kappa coefficient ranges from 23 % to about 2 times for CIrrMap250 compared to other irrigation maps. The overall accuracy of CIrrMap250 is 0.79, while it ranges from 0.60 to 0.73 for the other maps. The overall accuracy of CIrrMap250 improves from 8 % to 32 % compared to other maps. The producer's accuracy of irrigated samples is 0.75 in CIrrMap250, apparently higher than the other maps ranging from 0.40 to 0.64. The producer's accuracy of non-irrigated samples in CIrrMap250 is higher than higher than that of Zhu-map and Meier-map, but lower than EVI-map, NDVI-map, GI-map, and GIAM2000. The overall accuracy of all maps tends to be lower in arid and semi-arid areas than in humid and semi-humid areas. This is because non-irrigated cropland is more difficult to identify in arid and semi-arid areas, as indicated by the lower producer's accuracy of non-irrigation, although the contrary is the case for irrigated cropland. Regardless, the results indicate that CIrrMap250 has a higher accuracy

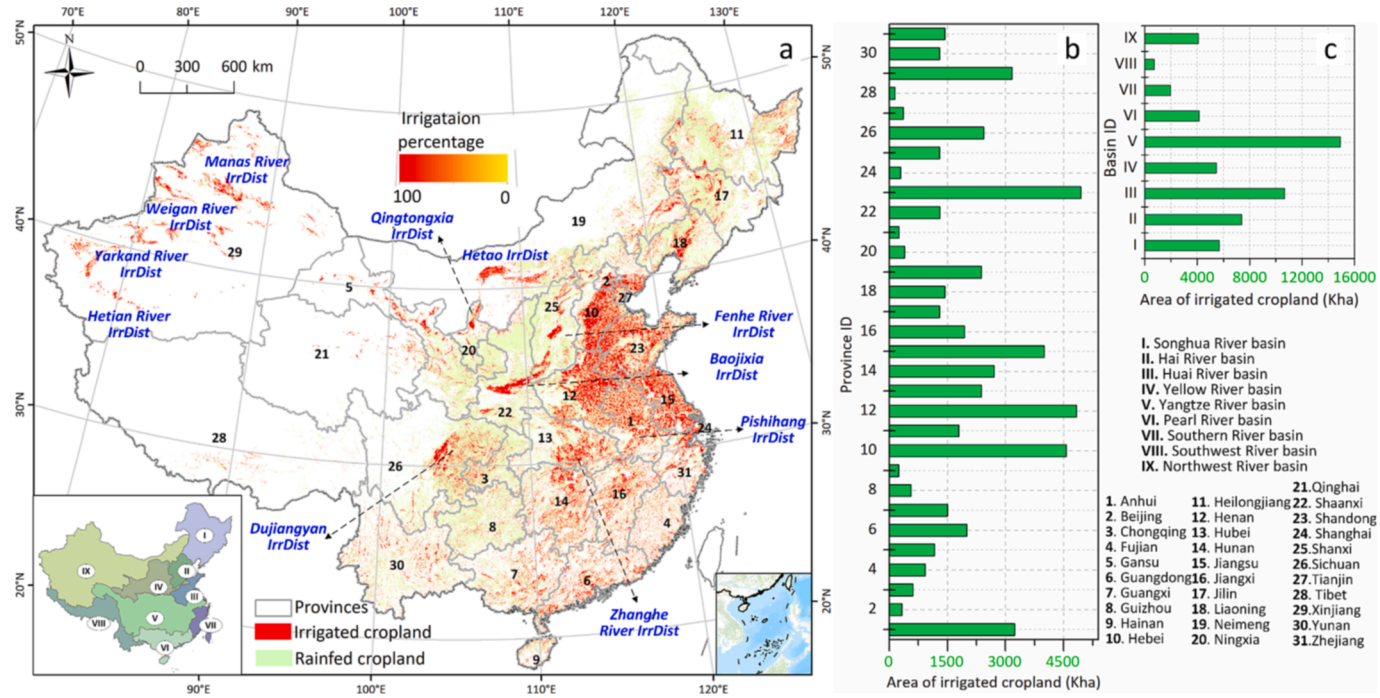

Fig. 3. Spatial distribution of irrigated and rainfed cropland in CIrrMap250. Blue words indicate some large-scale irrigation districts in mainland China. Area of irrigated croplands in different provinces and basins are summarized in panels b and c, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

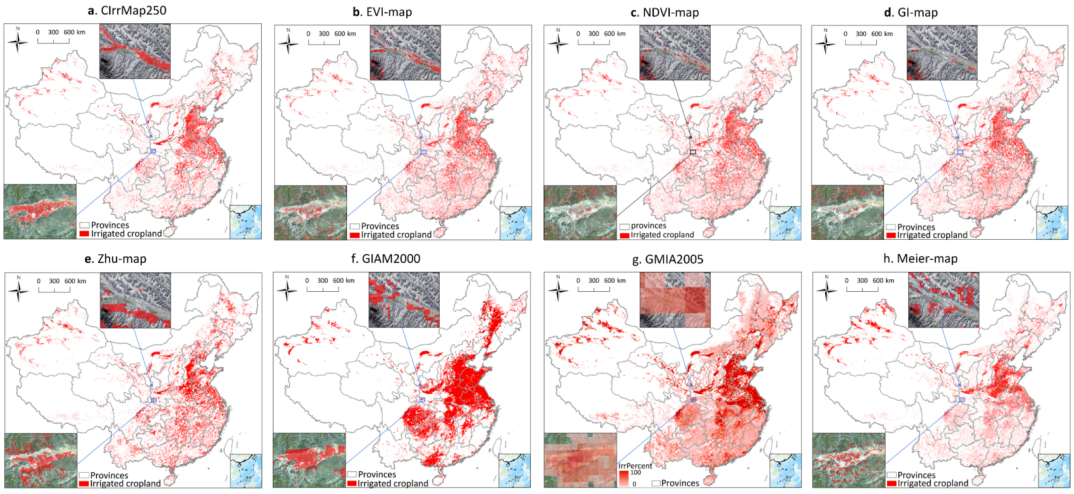

Fig. 4. Comparison of the spatial distribution of irrigated cropland in CIrrMap250 with other irrigation maps. a, CIrrMap250; b, EVI-map; c, NDVI-map; d, GI-map; e, Zhu-map; f, GIAM2000; g, GIMA2005; h, Meier-map. Inserted panels show the irrigated cropland distribution in two field-surveyed irrigation districts located in Gansu and Shaanxi provinces, respectively.

than the other irrigation maps.

4.4. Comparison of irrigated area estimates with statistics

Fig. 6 compares the irrigated area estimates of CIrrMap250, Zhu-map, GMIA2005, and Meier-map with the statistical data for the 3,652 mapping units of mainland China. GIAM2000 is not analyzed here due to the lack of pixel-level irrigation percentages. The irrigated area of CIrrMap250 shows a good agreement with the statistics, with the coefficient of determination (R^2) of 0.99, mainly due to the usage of irrigated area statistics in the generation of the training pool. The irrigated area of Zhu-map matches well with the provincial statistics, but the agreement is relatively low at the mapping unit level. This is because less than 50 % of the provinces in Zhu-map includes county-level irrigated area statistics (Zhu et al., 2014). For example, as shown in Fig. 8**b**, the irrigated areas in two mapping units of Heilongjiang Province are clearly overestimated because only the total irrigated area statistics in that province were used in the mapping process. The estimated irrigated area from GMIA2005 is generally consistent with the statistics, as the statistics were also used to generate GMIA2005, although not in as much detail as the data in our study. Among the four irrigation maps, Meier-map has the lowest agreement between the irrigated area estimates and the statistics because it combined statistical irrigated area with other remote sensing-detected irrigated area during the mapping process (Meier et al., 2018).

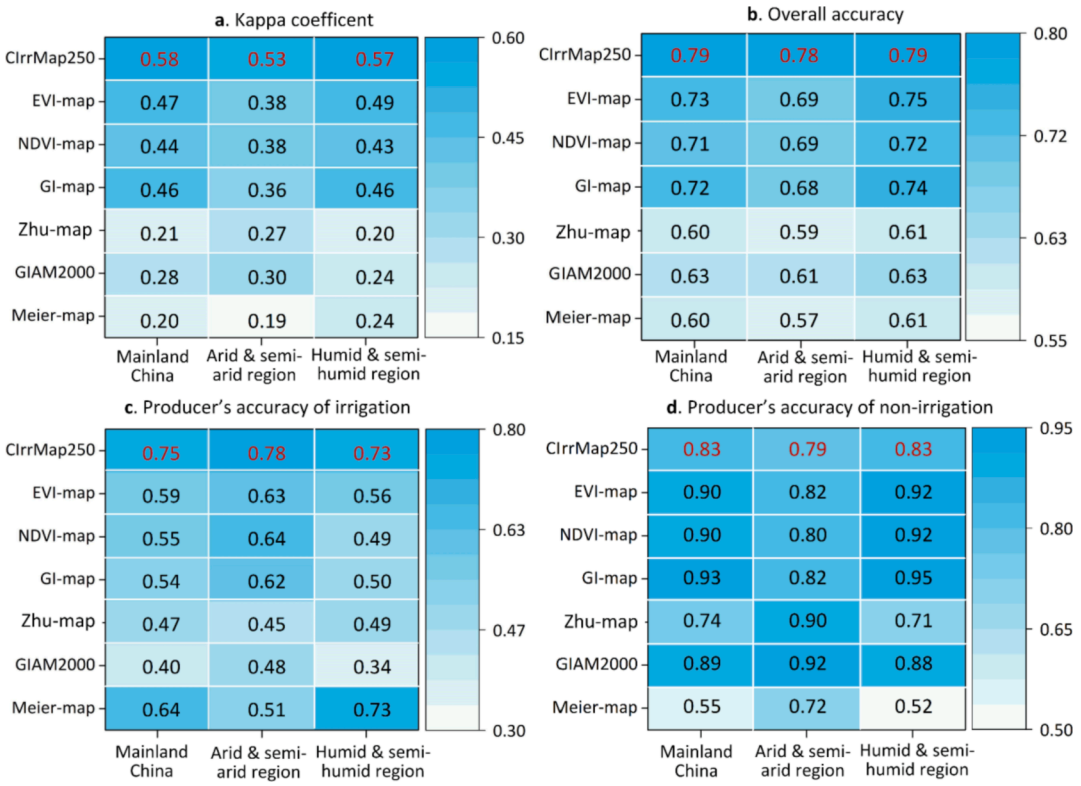

Fig. 5. Comparison of the Kappa coefficient (a), overall accuracy (b), and producer's accuracy (c, d) achieved by different irrigation maps. Results are shown for mainland China and its two subregions, i.e., the arid and semi-arid region (aridity index ≤ 0.5) and humid and semi-humid region (aridity index > 0.5).

5. Discussion

5.1. Irrigated cropland in different geographical feature ranges

We further calculated the proportion of irrigated cropland for seven irrigation maps at different elevations, slopes, and distances from water bodies. The North China Plain was excluded from the analysis due to its small elevation and slope gradients and the relatively consistent pattern of irrigated cropland in these maps. As shown in Fig. 7, CIrrMap250 has>70 % of irrigated cropland that is distributed in areas with elevation less than 500 m, slope less than 4°, and distance less than 3,000 m from water bodies. In comparison to other maps, more irrigated cropland in CIrrMap250 is located at lower elevations, on smaller slopes, and closer to water bodies. This is easy to understand because we incorporated irrigation suitability estimates based on elevation, slope, distance to water bodies, and aridity index into the irrigation mapping framework. As described in Section 3.3.1, areas with low elevation, gentle slope, small distance to water bodies are more favorable for irrigation. From this perspective, the spatial distribution seems to be more reasonable in CIrrMap250 than in other maps. Nevertheless, China has quite specific agriculture (e.g., terraces and smallholder cropping systems) that requires reliable reference irrigation maps to robustly assess the above results, especially at the regional scale. The fact that the other maps have more irrigated cropland at higher elevations, larger slopes, and farther away from water bodies than CIrrMap250 can be explained by their neglect of the mixed grid effects. China has many scattered, small-scale croplands that are mixed with forests/grasslands in mountainous areas. Vegetation in mixed cropland-forest/grassland areas may be "greener" than croplands located in river valleys, and they can easily be detected as irrigated cropland if the vegetation index is mainly considered for extracting irrigated area, as illustrated by the two field-surveyed irrigated districts.

5.2. Large uncertainty in irrigation ratio estimates

The sources of cropland area in China can be divided into four categories, including remote sensing, the National Bureau of Statistics of China (NBSC), the Ministry of Natural Resources of the People's Republic of China (MNRPRC), and the National Land Survey. The National Land Survey of China was conducted three times in 1984–1987, 2007–2009, and 2017–2019; while no data was available around 2000. We estimated irrigation ratio, i.e., the ratio of irrigated cropland area to total cropland area, for difference provinces in mainland China based on the cropland area from remote sensing (Xu et al., 2018), NBSC (National Bureau of Statistics of China, 2001a), and MNRPRC (Ministry of Land and Resources of the People's Republic of China, 2001). As shown in Fig. 8, irrigation ratio varies significantly among the 31 provinces, with higher values in Beijing, Shanghai, Jiangsu, and Xinjiang, while lower values in Heilongjiang, Yunnan, and Guizhou. Irrigation ratio also exhibits a large difference for the same region. For instance, it ranges from 0.54 to 0.93 in Xinjiang using cropland area from different sources. Irrigation ratio estimates based on the remote sensing-estimated cropland area are lower than the estimates based on the NBSC-reported cropland area and MNRPRC-reported cropland area. This is because, as shown in Fig. 8, the cropland acreages reported by NBSC and MNRPRC are obviously smaller than those estimated by remote sensing, although they exhibit a higher correlation ($R^2 > 0.85$). In this study, it is challenging to assess which sources of cropland area are reliable. Nevertheless, previous study indicates that remote sensing tends to overestimate cropland area due to the widespread presence of non-cropland within the map grid and the classification accuracy problems (Guo, 2006). On the other hand, NBSC tends to underestimate the cropland area due to administrative and technical reasons (Xu et al., 2014). Overall, irrigation ratio estimates have large uncertainties due to the varying cropland area from multiple sources.

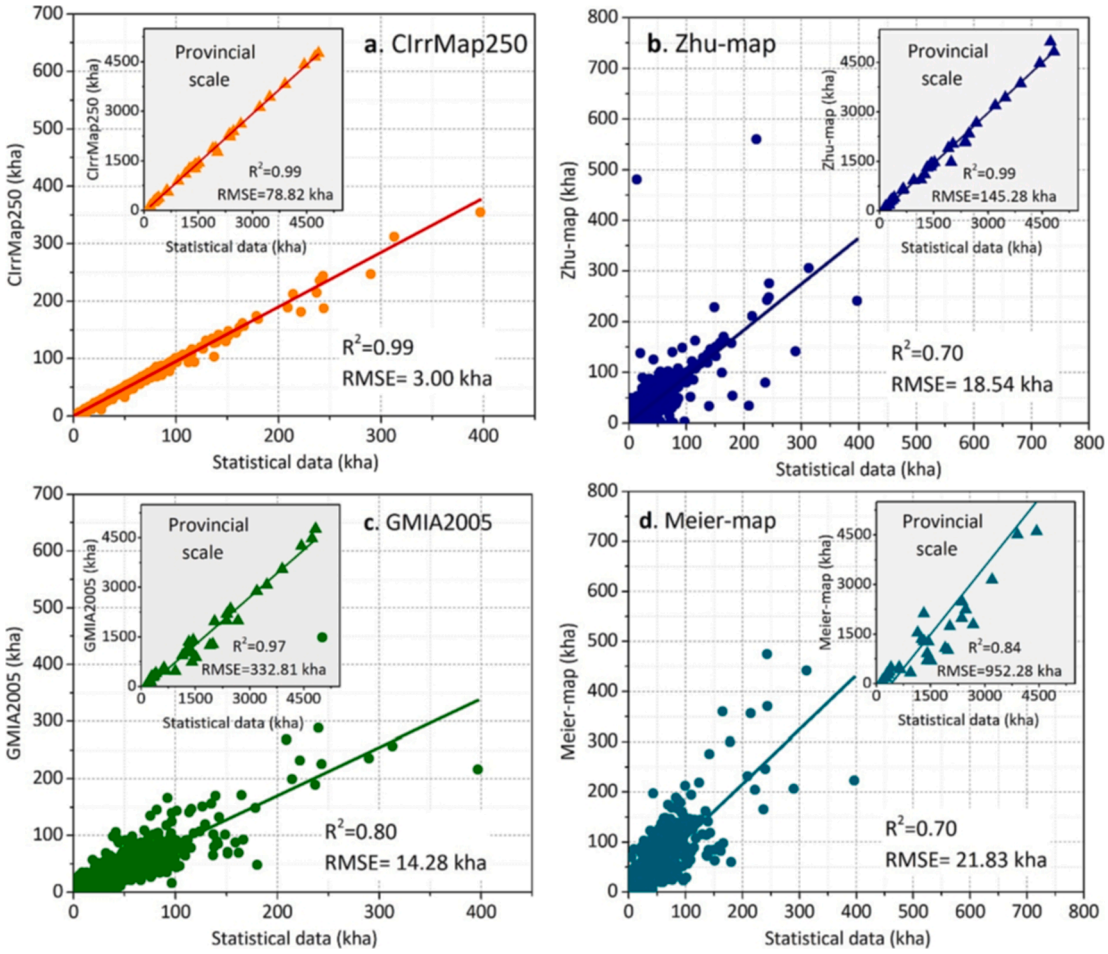

Fig. 6. Comparison of the irrigated area estimates against the statistics for 3,652 mapping units of mainland China. a, CIrrMap250; b, Zhu-map; c, GMIA2005; d, Meier-map. The inserted gray panels show the results at the provincial level.

5.3. Limitations and prospects

We admit that this study has some limitations and uncertainties that deserve further investigations in the future. To begin with, the proposed irrigation mapping method depends heavily on irrigated area statistics. Although the statistical irrigated areas were collected and reconstructed with an unprecedented level of detail in our study, they inevitably have some uncertainties and inconsistences (Ajaz et al., 2019). Furthermore, the statistics represent the effective irrigated area, i.e., the area equipped for irrigation, rather than the actual irrigated area, which may lead to biased results. In fact, the China Water Statistical Yearbook (CWSY) (China Ministry of Water Resources, 2001) have also reported the actual irrigated area at the provincial level, which is closely matched to the effective irrigated area with a coefficient of determination of 0.99 (Supplementary **Fig. S5)**. However, actual irrigated area at the city/county level is not available, which limits its use to constrain the irrigated cropland distribution in spatial detail. We further conducted an experiment to assign the actual provincial irrigated area to the county level based on the statistical effective irrigated area and recreated the irrigated cropland map using the method proposed in this study. We found that the spatial pattern of irrigated croplands in the new map was very similar to CIrrMap250 (Supplementary **Fig. S6)**, but showed no improvement in accuracy.

Second, this study incorporates irrigation suitability into the irrigated cropland mapping framework, which greatly improves the performance of irrigation map. However, the irrigation suitability analysis is somewhat subjective due to the requirement of weight setting. If sufficient ground truth data are available in the future, they can be divided into two parts, one for automatic weight calibration and one for performance evaluation. Meanwhile, due to the large number of mapping units (>3500), the irrigation suitability analysis was performed only at the provincial level, which may affect the accuracy of the irrigation map, especially in provinces with high topographic and climatic gradients, such as Gansu and Heilongjiang provinces. In addition, terraces are widely distributed in China, accounting for 26 % of the country's cropland area (Cao et al., 2021), and about 30 % of them are located in areas with slopes higher than 8° (Supplementary **Fig. S7**). There may be many irrigable terraces in high-slope areas, thus our assumption that gentle-slope cropland is more likely to be irrigated than steep-slope cropland may not hold true, adding some uncertainties to our results. Nevertheless, irrigation suitability is determined not only by slope but also by three other factors (i.e., elevation, DistToWater, and drought index), and meanwhile, for provinces with concentrated distribution of terraces, the weight of slope (less than0.16) was significantly lower than other factors in the irrigation suitability analysis, suggesting that our results may not be significantly influenced.

Lastly, this study only mapped irrigated cropland in mainland China for a specific year (i.e., 2000) to verify the reliability of the newly proposed mapping method, as done in many other studies (Ozdogan and Gutman, 2008; Xiang et al., 2019; Xie et al., 2019; Zhu et al., 2021; Zhang et al., 2022). The number of reference points is relatively small and some of them may not be representative of the actual situation in 2000, which may bring uncertainty to our results. The lack of sufficient ground truth data would be addressed in the future by conducting more field work and by collaborating with more researchers in the field of irrigated cropland mapping and with the Ministry of Natural Resources

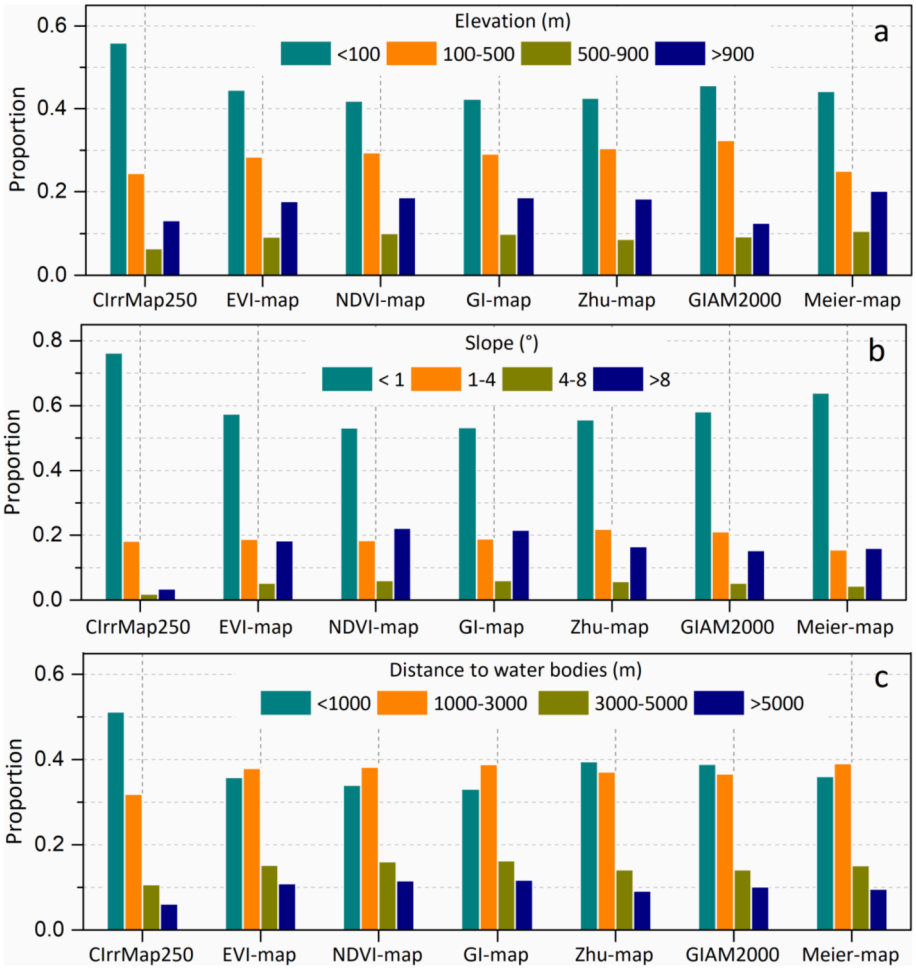

Fig. 7. Proportion of irrigated croplands over different ranges of elevation (a), slope (b), and distance to water bodies (c) in the seven irrigation maps.

of the People's Republic of China (https://www.mnr.gov.cn/), which has much reliable survey data. In the next step, we will apply our new method to generate time-continuous maps of irrigated cropland from 2000 to the present and collect more reference points and survey data spanning multiple years to robustly evaluate the performance of our new mapping method and the accuracy of irrigation maps.

6. Conclusions

This study proposed a novel approach to develop the first 250-m irrigated cropland map in mainland China (CIrrMap250) for the year 2000 by integrating remote sensing, irrigation suitability, and irrigated area statistics. The novelty of the new approach lies in the incorporation of irrigation suitability into the irrigation mapping process. The new method consists of three major steps, including irrigation suitability analysis, training pool generation based on suitability-adjusted vegetation indices and irrigated area statistics, and machine learning-based classification of irrigated and rainfed cropland. We evaluated the performance of CIrrMap250 and compared it with three irrigation maps (i.e., EVI-map, NDVI-map, and GI-map) generated using the threshold-based classification method, as well as four other existing maps including GMIA2005 (Siebert et al., 2005), GIAM2000 (Thenkabail et al., 2009), Zhu-map (Zhu et al., 2014), and Meier-map (Meier et al., 2018). The accuracy of the irrigation maps was assessed from three different aspects, including the distribution of irrigated croplands, agreement with reference points, and consistency with irrigated area statistics.

Results indicate that the irrigated croplands are mainly distributed in the North China Plain, Northwest China, and some large-scale irrigation districts. Shandong province holds the largest irrigated area, followed by Henan, Hebei, Jiangsu, Anhui, and Xinjiang, while at the basin scale, the Yangtze River basin has the largest irrigated area, followed by the Huai River basin, Hai River basin, and Yellow River basin. Intensively irrigated areas are overall well captured by CIrrMap250 and all the other maps. CIrrMap250 has a similar irrigation distribution pattern to EVI-map, NDVI-map, due to the consistent constraints of irrigated area statistics on the mapping results. However, there are notable differences between CIrrMap250 and the four existing maps at both regional and local scales. All maps except CIrrMap250 tend to underestimate irrigated cropland in river valleys and overestimate irrigated cropland in the mountainous areas, as evidenced by the field-surveyed irrigation districts, mainly because of their neglect of the mixed grid effects. CIrrMap250 has a higher point-scale accuracy than the other maps. CIrrMap250 achieves an improvement in Kappa coefficient by 23 % to about 2 times and an improvement in overall accuracy by 8 % to 32 %, in comparison to other irrigation maps. Additionally, the irrigated area of CIrrMap250 shows a good agreement with the statistics, mainly due to the utilization of irrigated area statistics in the generation of the training pool.

The irrigation mapping method can be easily applied to other regions of the world, and the new irrigated cropland map has great potential to support agriculture, climate, and environment research as well as water resource management in mainland China. Nevertheless, we admit that this study has some limitations and uncertainties, such as the heavy dependence on irrigated area statistics and the subjectivity of irrigation suitability analysis, which deserve further investigations in the future. In

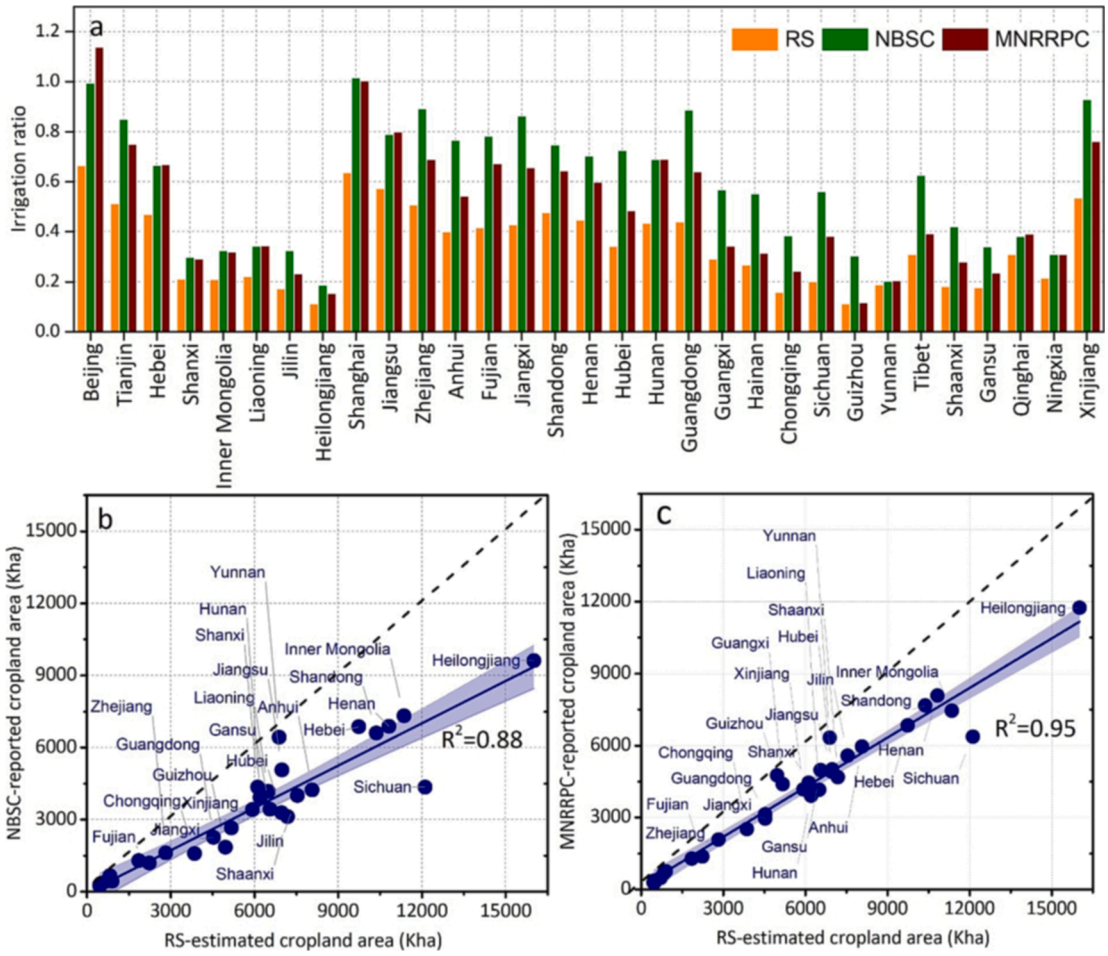

Fig. 8. Irrigation ratio estimates for different provinces based on cropland area from different sources. RS, NBAC, and MNRRPC denote remote sensing, National Bureau of Statistics of China, and Ministry of Natural Resources of the People's Republic of China, respectively. Panels b and c show scatter plots of the cropland area estimated by RS versus the area reported by NBSC and MNRRPC, respectively.

the next step, we will attempt to generate time-continuous maps of irrigated cropland from 2000 to the present and collect more reference points and survey data to robustly evaluate the performance of our new mapping method and the accuracy of irrigation maps.

7. Data and code availability

CIrrMap250 can be accessed at: https://doi.org/10.6084/m9.figshare.17056442.v2. All codes for this study are freely available at via: https://github.com/HydroRS/CIrrMap.

CRediT authorship contribution statement

Ling Zhang: Conceptualization, Methodology, Funding acquisition, Investigation, Software, Resources, Visualization, Writing – original draft, Writing – review & editing. **Kun Zhang:** Investigation, Resources, Visualization, Writing – review & editing. **Xiufang Zhu:** Data curation, Formal analysis, Writing – review & editing. **Hao Chen:** Data curation, Writing – review & editing. **Weizhen Wang:** Funding acquisition, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was supported by the major project of China high-resolution earth observation system (21-Y20B01-9001-19/22), the National Natural Science Foundation of China (41901045), the CAS "Light of West China" Program, and the Opening Research Foundation of Key Laboratory of Land Surface Process and Climate Change in Cold and Arid Regions, Chinese Academy of Sciences (LPCC2020004).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2022.128413.

References

Ajaz, A., Karimi, P., Cai, X., De Fraiture, C., Akhter, M.S., 2019. Statistical Data Collection Methodologies of Irrigated Areas and Their Limitations: A Review. Irrig. Drain. 68 (4), 702–713.
Akıncı, H., Özalp, A.Y., Turgut, B., 2013. Agricultural land use suitability analysis using GIS and AHP technique. Comput. Electron. Agric. 97, 71–82.
Ambika, A.K., Wardlow, B., Mishra, V., 2016a. Remotely sensed high resolution irrigated area mapping in India for 2000 to 2015. Sci. Data 3 (1), 160118.
Ambika, A.K., Wardlow, B., Mishra, V., 2016b. Remotely sensed high resolution irrigated area mapping in India for 2000 to 2015. Sci. Data 3 (1).
Bazzi, H., Baghdadi, N., Amin, G., Fayad, I., Zribi, M., Demarez, V., Belhouchette, H., 2021. An Operational Framework for Mapping Irrigated Areas at Plot Scale Using Sentinel-1 and Sentinel-2 Data. Remote Sensing 13 (13), 2584.
Biggs, T.W., Thenkabail, P.S., Gumma, M.K., Scott, C.A., Parthasaradhi, G.R., Turral, H. N., 2006. Irrigated area mapping in heterogeneous landscapes with MODIS time

series, ground truth and census data, Krishna Basin, India. Int. J. Remote Sensing 27 (19), 4245–4266.
Breiman, L., 2001. Random Forests. Machine Learning 45 (1), 5–32.
Cao, B., Yu, L., Naipal, V., Ciais, P., Li, W., Zhao, Y., Wei, W., Chen, D., Liu, Z., Gong, P., 2021. A 30 m terrace mapping in China using Landsat 8 imagery and digital elevation model based on the Google Earth Engine. Earth Syst. Sci. Data 13 (5), 2437–2456.
Chen, Y., Yu, J., Khan, S., 2010. Spatial sensitivity analysis of multi-criteria weights in GIS-based land suitability evaluation. Environ. Modell. Software 25 (12), 1582–1591.
Chen, Y., Lu, D., Luo, L., Pokhrel, Y., Deb, K., Huang, J., Ran, Y., 2018. Detecting irrigation extent, frequency, and timing in a heterogeneous arid agricultural region using MODIS time series, Landsat imagery, and ancillary data. Remote Sens. Environ. 204, 197–211.
China Ministry of Water Resources, 2001. China Water Statistical Yearbook (CWSY) China Water&Power Press.
Cuo, L., Zhang, Y., Gao, Y., Hao, Z., Cairang, L., 2013. The impacts of climate change and land cover/use transition on the hydrology in the upper Yellow River Basin, China. J. Hydrol. 502, 37–52.
Dari, J., Quintana-Seguí, P., José Escorihuela, M., Stefan, V., Brocca, L., Morbidelli, R., 2021. Detecting and mapping irrigated areas in a Mediterranean environment by using remote sensing soil moisture and a land surface model. J. Hydrol. 126129.
Debeurs, K., Townsend, P., 2008. Estimating the effect of gypsy moth defoliation using MODIS. Remote Sens. Environ. 112 (10), 3983–3990.
Deines, J.M., Kendall, A.D., Hyndman, D.W., 2017. Annual Irrigation Dynamics in the U. S. Northern High Plains Derived from Landsat Satellite Data. Geophys. Res. Lett. 44 (18), 9350–9360.
Deines, J.M., Kendall, A.D., Crowley, M.A., Rapp, J., Cardille, J.A., Hyndman, D.W., 2019. Mapping three decades of annual irrigation across the US High Plains Aquifer using Landsat and Google Earth Engine. Remote Sens. Environ. 233, 111400.
Dong, T., Zuo, L., Zhang, Z., 2009. Extraction of irrigated croplans from MODIS (In Chinese). Journal of Remote Sensing 13 (3), 528–534.
Gao, B.-C., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58 (3), 257–266.
Gao, Q., Zribi, M., Escorihuela, M., Baghdadi, N., Segui, P., 2018. Irrigation Mapping Using Sentinel-1 Time Series at Field Scale. Remote Sensing 10 (9), 1495.
Gardner, M.W., Dorling, S.R., 1998. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmos. Environ. 32 (14), 2627–2636.
Gitelson, A.A., 2005. Remote estimation of canopy chlorophyll content in crops. Geophys. Res. Lett. 32 (8).
Guo, G., 2006. How much cropland in China-Dicussion on the accuaracy of remote sensing (in Chinese). J. Inner Mongolia Agric. Univ. 27 (2), 56–59.
He, J., Yang, K., 2016. China meteorological forcing dataset (1979-2015). National Tibetan Plateau Data Center.
He, J., Yang, K., Tang, W., Lu, H., Qin, J., Chen, Y., Li, X., 2020. The first high-resolution meteorological forcing dataset for land process studies over China. Sci. Data 7 (1).
Hilker, T., Lyapustin, A.I., Tucker, C.J., Sellers, P.J., Hall, F.G., Wang, Y., 2012. Remote sensing of tropical ecosystems: Atmospheric correction and cloud masking matter. Remote Sens. Environ. 127, 370–384.
Huang, G., Zhou, H., Ding, X., Zhang, R., 2012. Extreme Learning Machine for Regression and Multiclass Classification. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 42 (2), 513–529.
Huete, A.R., Liu, H.Q., Batchily, K., van Leeuwen, W., 1997. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. Remote Sens. Environ. 59 (3), 440–451.
IDCD, 2018. International Commission on Irrigation and Drainage, World Irrigated Area.
Ishikawa, Y., Yamazaki, D., 2021. Global high-resolution estimation of cropland suitability and its comparative analysis to actual cropland distribution. Hydrological Research Letters 15 (2), 9–15.
Karthikeyan, L., Chawla, I., Mishra, A.K., 2020. A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses. J. Hydrol. 586, 124905.
Kumar, S.V., Peters-Lidard, C.D., Santanello, J.A., Reichle, R.H., Draper, C.S., Koster, R. D., Nearing, G., Jasinski, M.F., 2015. Evaluating the utility of satellite soil moisture retrievals over irrigated areas and the ability of land data assimilation methods to correct for unmodeled processes. Hydrol. Earth Syst. Sci. 19 (11), 4463–4478.
Li, H., Chen, Y., 2020. Assessing potential land suitable for surface irrigation using groundwater data and multi-criteria evaluation in Xinjiang inland river basin. Comput. Electron. Agric. 168, 105079.
Li, Y., Yang, X., Cai, H., Xiao, L., Xu, X., Liu, L., 2015. Topographical Characteristics of Agricultural Potential Productivity during Cropland Transformation in China. Sustainability 7 (1), 96–110.
Liu, Y., Li, Q., Wu, W., 2022. Analysis of feature slection for mapping irrigated cropland in northern China (In Chinese). Chin. J. Agric. Resour. Regional Plann. 42 (9), 27–35.
Liu, J., Liu, M., Tian, H., Zhuang, D., Zhang, Z., Zhang, W., Tang, X., Deng, X., 2005. Spatial and temporal patterns of China's cropland during 1990–2000: An analysis based on Landsat TM data. Remote Sens. Environ. 98 (4), 442–456.
Liu, Y., Wu, W., Li, H., Imtiaz, M., Li, Z., Zhou, Q., 2018. Intercomparison on Four Irrigated Cropland Maps in Mainland China. Sensors 18 (4), 1197.
Lu, Y., Song, W., Lü, J., Chen, M., Su, Z., Zhang, X., Li, H., 2021. A pixel-based spectral matching method for mapping high-resolution irrigated areas using EVI time series. Remote Sens. Lett. 12 (2), 169–178.
Mandal, B., Dolui, G., Satpathy, S., 2018. Land suitability assessment for potential surface irrigation of river catchment for irrigation development in Kansai watershed, Purulia, West Bengal. India. Sustainable Water Resources Management 4 (4), 699–714.
Massari, C., Modanesi, S., Dari, J., Gruber, A., De Lannoy, G.J.M., Girotto, M., Quintana-Seguí, P., Le Page, M., Jarlan, L., Zribi, M., Ouaadi, N., Vreugdenhil, M., Zappa, L., Dorigo, W., Wagner, W., Brombacher, J., Pelgrum, H., Jaquot, P., Freeman, V., Volden, E., Fernandez Prieto, D., Tarpanelli, A., Barbetta, S., Brocca, L., 2021. A Review of Irrigation Information Retrievals from Space and Their Utility for Users. Remote Sensing 13 (20), 4112.
McDermid, S.S., Mahmood, R., Hayes, M.J., Bell, J.E., Lieberman, Z., 2021. Minimizing trade-offs for sustainable irrigation. Nat. Geosci. 14 (10), 706–709.
Meier, J., Zabel, F., Mauser, W., 2018. A global approach to estimate irrigated areas – a comparison between different data and statistics. Hydrol. Earth Syst. Sci. 22 (2), 1119–1133.
Ministry of Land and Resources of the People's Republic of China (Ed.), 2001. China Land & Resources Almanac (in Chinese). https://data.cnki.net/yearbook/Single/N2013100006.
Nagaraj, D., Proust, E., Todeschini, A., Rulli, M.C., D'Odorico, P., 2021. A new dataset of global irrigation areas from 2001 to 2015. Adv. Water Resour. 152, 103910.
National Bureau of Statistics of China, 2001a. China Statistical Yearbook (in Chinese). China Statistics Press. http://www.stats.gov.cn/tjsj/ndsj/2001c/mulu.htm.
National Bureau of Statistics of China, 2001b. China Statistical Yearbook for Regional Economy (in Chinese). China Finance and Economy Press 1–504.
Ozdogan, M., Gutman, G., 2008. A new methodology to map irrigated areas using multi-temporal MODIS and ancillary data: An application example in the continental US. Remote Sens. Environ. 112 (9), 3520–3537.
Ozdogan, M., Yang, Y., Allez, G., Cervantes, C., 2010. Remote Sensing of Irrigated Agriculture: Opportunities and Challenges. Remote Sensing 2 (9), 2274–2304.
Pageot, Y., Baup, F., Inglada, J., Baghdadi, N., Demarez, V., 2020. Detection of Irrigated and Rainfed Crops in Temperate Areas Using Sentinel-1 and Sentinel-2 Time Series. Remote Sensing 12 (18), 3044.
Peña-Arancibia, J.L., McVicar, T.R., Paydar, Z., Li, L., Guerschman, J.P., Donohue, R.J., Dutta, D., Podger, G.M., van Dijk, A.I.J.M., Chiew, F.H.S., 2014. Dynamic identification of summer cropping irrigated areas in a large basin experiencing extreme climatic variability. Remote Sens. Environ. 154, 139–152.
Pervez, M.S., Brown, J.F., 2010. Mapping Irrigated Lands at 250-m Scale by Merging MODIS Data and National Agricultural Statistics. Remote Sensing 2 (10), 2388–2412.
Portmann, F.T., Siebert, S., Döll, P., 2010. MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. Global Biogeochem. Cycles 24 (1), n/a-n/a.
Priestley, C.H.B., Taylor, R.J., 1972. On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. Mon. Weather Rev. 100 (2), 81–92.
Puy, A., Borgonovo, E., Lo Piano, S., Levin, S.A., Saltelli, A., 2021. Irrigated areas drive irrigation water withdrawals. Nat. Commun. 12 (1).
Rosa, L., Chiarelli, D.D., Sangiorgio, M., Beltran-Peña, A.A., Rulli, M.C., D Odorico, P., Fung, I., 2020. Potential for sustainable irrigation expansion in a 3 °C warmer climate. Proceedings of the National Academy of Sciences: 202017796.
Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. In: Proc. Third Earth Resources Technology Satellite-1 Symposium, SP-351, Greenbelt, MD, pp. 309–317.
Salmon, J.M., Friedl, M.A., Frolking, S., Wisser, D., Douglas, E.M., 2015. Global rain-fed, irrigated, and paddy croplands: A new high resolution map derived from remote sensing, crop inventories and climate data. Int. J. Appl. Earth Obs. Geoinf. 38, 321–334.
Shahriar Pervez, M., Budde, M., Rowland, J., 2014. Mapping irrigated areas in Afghanistan over the past decade using MODIS NDVI. Remote Sens. Environ. 149, 155–165.
Sharma, A.K., Hubert-Moy, L., Buvaneshwari, S., Sekhar, M., Ruiz, L., Moger, H., Bandyopadhyay, S., Corgne, S., 2021. Identifying Seasonal Groundwater-Irrigated Cropland Using Multi-Source NDVI Time-Series Images. Remote Sensing 13 (10), 1960.
Siebert, S., Döll, P., Hoogeveen, J., Faures, J.M., Frenken, K., Feick, S., 2005. Development and validation of the global map of irrigation areas. Hydrol. Earth Syst. Sci. 9 (5), 535–547.
Siebert, S., Döll, P., 2010. Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation. J. Hydrol. 384 (3–4), 198–217.
Thenkabail, P.S., Biradar, C.M., Noojipady, P., Dheeravath, V., Li, Y., Velpuri, M., Gumma, M., Gangalakunta, O.R.P., Turral, H., Cai, X., Vithanage, J., Schull, M.A., Dutta, R., 2009. Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium. Int. J. Remote Sens. 30 (14), 3679–3733.
Wada, Y., Wisser, D., Eisner, S., Flörke, M., Gerten, D., Haddeland, I., Hanasaki, N., Masaki, Y., Portmann, F.T., Stacke, T., Tessler, Z., Schewe, J., 2013. Multimodel projections and uncertainties of irrigation water demand under climate change. Geophys. Res. Lett. 40 (17), 4626–4632.
Wang, X., Müller, C., Elliot, J., Mueller, N.D., Ciais, P., Jägermeyr, J., Gerber, J., Dumas, P., Wang, C., Yang, H., Li, L., Deryng, D., Folberth, C., Liu, W., Makowski, D., Olin, S., Pugh, T.A.M., Reddy, A., Schmid, E., Jeong, S., Zhou, F., Piao, S., 2021. Global irrigation contribution to wheat and maize yield. Nat. Commun. 12 (1).
Worqlul, A.W., Collick, A.S., Rossiter, D.G., Langan, S., Steenhuis, T.S., 2015. Assessment of surface water irrigation potential in the Ethiopian highlands: The Lake Tana Basin. Catena 129, 76–85.
Worqlul, A.W., Jeong, J., Dile, Y.T., Osorio, J., Schmitter, P., Gerik, T., Srinivasan, R., Clark, N., 2017. Assessing potential land suitable for surface irrigation using groundwater in Ethiopia. Appl. Geogr. 85, 1–13.

Xiang, K., Ma, M., Liu, W., Dong, J., Zhu, X., Yuan, W., 2019. Mapping Irrigated Areas of Northeast China in Comparison to Natural Vegetation. Remote Sensing 11 (7), 825.
Xiang, K., Yuan, W., Wang, L., Deng, Y., 2020. An LSWI-Based Method for Mapping Irrigated Areas in China Using Moderate-Resolution Satellite Data. Remote Sensing 12 (24), 4181.
Xie, Y., Lark, T.J., Brown, J.F., Gibbs, H.K., 2019. Mapping irrigated cropland extent across the conterminous United States at 30 m resolution using a semi-automatic training approach on Google Earth Engine. ISPRS J. Photogramm. Remote Sens. 155, 136–149.
Xie, Y., Gibbs, H.K., Lark, T.J., 2021. Landsat-based Irrigation Dataset (LANID): 30-m resolution maps of irrigation distribution, frequency, and change for the U.S., 1997–2017. Earth Syst. Sci. Data Discuss. 2021, 1–32.
Xie, Y., Lark, T.J., 2021. Mapping annual irrigation from Landsat imagery and environmental variables across the conterminous United States. Remote Sens. Environ. 260, 112445.
Xu, J., Yang, H., Huang, X., 2014. Differences in statistical arable land area based on different data sources and related research (in Chinese). Proceedings of the 2014 Annual Academic Conference of the Land Institute of China, 438-444. DOI: 10.26914/c.cnkihy.2014.002217.
Xu, X., Liu, J., Zhang, S., Li, R., Yan, C., Wu, S., 2018. Remote sensing-based monitoring dataset of land use and cover change over multiple periods in China (CNLUCC) (in Chinese). Resource and Environmental Science Data Center. https://doi.org/10.12078/2018070201.
Yin, L., Feng, X., Fu, B., Chen, Y., Wang, X., Tao, F., 2020. Irrigation water consumption of irrigated cropland and its dominant factor in China from 1982 to 2015. Adv. Water Resour. 143, 103661.
Zaussinger, F., Dorigo, W., Gruber, A., Tarpanelli, A., Filippucci, P., Brocca, L., 2019. Estimating irrigation water use over the contiguous United States by combining satellite and reanalysis soil moisture data. Hydrol. Earth Syst. Sci. 23 (2), 897–923.
Zhang, C., Dong, J., Xie, Y., Zhang, X., Ge, Q., 2022. Mapping irrigated croplands in China using a synergetic training sample generating method, machine learning classifier, and Google Earth Engine. Int. J. Appl. Earth Obs. Geoinf. 112, 102888.
Zhang, L., Li, X., Zheng, D., Zhang, K., Ma, Q., Zhao, Y., Ge, Y., 2021. Merging multiple satellite-based precipitation products and gauge observations using a novel double machine learning approach. J. Hydrol. 125969.
Zhu, Z., Zhang, Z., Zuo, L., Sun, F., Pan, T., Li, J., Zhao, X., Wang, X., 2021. The Detecting of Irrigated Croplands Changes in 1987–2015 in Zhangjiakou. IEEE Access 9, 96076–96091.
Zhu, X., Zhu, W., Zhang, J., Pan, Y., 2014. Mapping Irrigated Areas in China From Remote Sensing and Statistical Data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7 (11), 4490–4504.
Zohaib, M., Kim, H., Choi, M., 2019. Detecting global irrigated areas by using satellite and reanalysis products. Sci. Total Environ. 677, 679–691.