Remote Sensing of Environment 260 (2021) 112445

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Mapping annual irrigation from Landsat imagery and environmental variables across the conterminous United States

Yanhua Xie [a,b,*], Tyler J. Lark [a,b]

[a] Nelson Institute Center for Sustainability and the Global Environment (SAGE), University of Wisconsin-Madison, 1710 University Avenue, Madison, WI 53726, USA
[b] DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI 53726, USA

ARTICLE INFO

Editor: Marie Weiss

Keywords:
Irrigation
Irrigation change
Water use
United States
Landsat
Google earth engine

ABSTRACT

Identifying the location of irrigated croplands and how they change over time is critical for assessing and managing limited water resources to navigate such challenges as local to global water scarcity, increasing demands for food and energy production, and environmental sustainability. Although efforts have been made to map irrigated area for the U.S., multi-year nationwide maps at field-relevant resolutions are still unavailable; existing products suffer from coarse resolution, uncertain accuracy, and/or limited spatial coverage, especially in the eastern U.S. In this study, we present an approach to map the extent of irrigated croplands across the conterminous U.S. (CONUS) for each year in the period of 1997–2017. To scale nationwide, we developed novel methods to generate training datasets covering both the western and eastern U.S. For the more arid western U.S., we built upon the methods of Xie et al. (2019) and further developed a greenness-based normalization technique to estimate optimal thresholds of crop greenness in any year based on those in USDA NASS census years (i.e., 1997, 2002, 2007, 2012, and 2017). For the relatively humid eastern states, we collected data on the current status of center pivot irrigated and non-irrigated fields and extended these sample points through time using various indices and observational thresholds. We used the generated samples along with remote sensing features and environmental variables to train county-stratified random forest classifiers annually for pixel-level classification of irrigated extent each year and subsequently implemented a logic-based post-classification filtering. The produced Landsat-based Irrigation Dataset (LANID-US) accurately reconstructed NASS irrigation patterns at both the county and state level while also supplying new annual area estimates for intra-epoch years. Nationwide pixel-level locational assessment further demonstrated an overall accuracy above 90% across years. In the 21-year study period, we found several hotspots of irrigation change including significant gains in the U.S. Midwest, the Mississippi River Alluvial Plain, and the East Coast as well as irrigation declines in the central and southern High Plains Aquifer and the southern California Central Valley, Arizona, and Florida. The resulting 30 m resolution LANID-US products represent the finest resolution account of nationwide irrigation use and dynamics across the United States to date. The developed approach, training data, and products are further extendable to other years (either before 1997 or after 2017) for continuous monitoring of irrigated area over CONUS and are spatially applicable to other regions with similar climate and cropping landscapes.

1. Introduction

Irrigation covers approximately 20% of worldwide croplands, supports 40% of global food production, and consumes about 70% of freshwater withdrawals (Rosegrant et al., 2009; Scanlon et al., 2012; Wada et al., 2013). Although irrigation accounts for a smaller proportion of total freshwater withdrawals in the U.S. (42%) (Dieter et al., 2018), there is increasing public concern about irrigation-induced impacts, such as groundwater depletion in the High Plains Aquifer and the California Central Valley (Brown and Pervez, 2014; Scanlon et al., 2012; Zektser et al., 2005), regional climate change in the Midwest and Great Plains (DeAngelis et al., 2010; Nocco et al., 2019), and groundwater pollution in central Minnesota and Wisconsin (Kraft and Stites, 2003; Shrestha et al., 2010). To fully evaluate the environmental impacts of irrigation and formulate strategies towards sustainable use of limited water resources, it is essential to know the precise location of irrigated

* Corresponding author at: Nelson Institute Center for Sustainability and the Global Environment (SAGE), University of Wisconsin-Madison, 1710 University Avenue, Madison, WI 53726, USA.
E-mail address: xie78@wisc.edu (Y. Xie).

https://doi.org/10.1016/j.rse.2021.112445
Received 30 November 2020; Received in revised form 4 April 2021; Accepted 9 April 2021

agriculture and how it changes over time. This need is further magnified by increasing demands for agricultural production under climate change and population growth (Lark et al., 2015; McDonald et al., 2011; Rosegrant et al., 2009; Seager et al., 2012; Seto et al., 2012).

Remote sensing has emerged as a powerful tool to produce a wide range of land use/cover products, but few explicitly include irrigation as a separate class while others that do are not well updated. Several global initiatives of irrigation mapping that cover the U.S. include the U.S. Geological Survey (USGS) Global Land Cover Map (Loveland et al., 2000), Global Map of Irrigation Areas (Siebert et al., 2005; Siebert et al., 2015), Global Irrigated Area Map (Thenkabail et al., 2009), Global Cropland Area Database (Teluguntla et al., 2015), and the Global Rainfed, Irrigated and Paddy Croplands dataset (Salmon et al., 2015). Several U.S.-specific nationwide irrigation maps are also available, including the Moderate Resolution Imaging Spectroradiometer (MODIS) based Irrigated Agriculture Dataset for the U.S. (MIrAD-US) (Brown and Pervez, 2014; Pervez and Brown, 2010), the MODIS-derived map generated by Ozdogan and Gutman (2008), and county-level irrigation census data provided by U.S. Department of Agriculture (USDA) National Agricultural Statistics Service (NASS). The spatial resolution of these maps, however, ranges from 250 m to kilometers to broad administrative units. This coarse spatial resolution is problematic for many localized applications due to spatial imprecision and uncertainty, inaccurate field edge characterization, and the frequent omission of fragmented fields (Brown and Pervez, 2014; Deines et al., 2017; Ozdogan and Gutman, 2008; Wardlow and Callahan, 2014; Xie et al., 2019). Furthermore, accuracies of these maps are often low, especially in the more humid eastern U.S., where high confusion between irrigated and rainfed croplands exists (Ozdogan and Gutman, 2008; Xie et al., 2019).

To address the limited spatial details of existing maps, Xie et al. (2019) created a 30 m Landsat-based Irrigation Dataset for the year 2012 (i.e., LANID-US 2012). However, as was the case with other products mapped in a single date (e.g., Ozdogan and Gutman, 2008) or five-year intervals (e.g., Brown and Pervez, 2014), the infrequent outputs limit the tracking of irrigation across space and time. In contrast to its mapping, irrigation decision making is extremely dynamic due to annual crop rotations, fallow practices, climate variation, as well as social and economic factors (Drysdale and Hendricks, 2018; Sampson et al., 2019; Tack et al., 2017). As such, the lack of timely, annual information on irrigated extent effectively precludes analysis of irrigation change and associated impacts, thereby limiting valuable insights for understanding water use and availability.

In recent years, the open access of remotely sensed images (particularly the Landsat archive), improvement of computing capacity (e.g., Google Earth Engine), and increasingly available ancillary datasets have enabled finer scale irrigation mapping in a more frequent manner. However, gaps remain in the production of annual, fine-scale irrigation maps for the U.S. For instance, while annual 30 m resolution maps are recently available for select regions (Deines et al., 2019; Deines et al., 2017; Ketchum et al., 2020; Xu et al., 2019), they focus mostly on the arid and semi-arid western U.S., where irrigation mapping is more straightforward and routine. Despite the importance of irrigation mapping in the western U.S. for managing scarce water resources, such data is also critical across the Midwest and Eastern U.S., where conversion to irrigated agriculture is accelerating and concern of irrigation-induced impacts on hydrology and water quality is increasing (Nocco et al., 2019; Shrestha et al., 2010). Providing information on nationwide irrigated extent and dynamics would thus benefit studies aiming to evaluate effects of irrigation across many diverse ecosystems and agricultural regions that lie outside the arid West (Kraft and Stites, 2003; Nocco et al., 2019; Vashisht et al., 2015).

The availability of reference data of sufficient quantify, distribution, and representation is also essential to remote sensing classification and accuracy assessment. This is especially true for irrigation mapping, where the differences between irrigation-induced biomass changes and those in rainfed croplands are subtle in satellite-derived indices, particularly in humid areas (Ozdogan and Gutman, 2008; Pageot et al., 2020; Xie et al., 2019). The difficulty is further magnified when training data is needed across large spatial extents such as the conterminous U.S. (CONUS) and at frequent time intervals (e.g., annual). To address a general lack of training data, global irrigation maps usually develop irrigated vs. non-irrigated references from multi-source data sets, such as agricultural census statistics, national reports, and other related products (Siebert et al., 2005; Teluguntla et al., 2015; Thenkabail et al., 2009). As expected, these inconsistent, low locational accuracy reference data result in irrigation maps with limited performance in detecting the distribution of irrigation at local to regional scales.

In contrast to large-area studies, where selecting globally consistent, high accuracy ground truth data is often prohibitively cost-intensive, collecting sufficient and accurate training samples for local to regional studies is more feasible and efficient. For example, Deines et al. (2019; 2017) and Ketchum et al. (2020) relied on extensive manually collected samples to map irrigation in the western CONUS over time. They also proved that feature extension might be practical in arid and semi-arid areas where a temporally generalized classifier calibrated from training samples of selected years could be applied over longer periods of time. However, given the variabilities in crop growth caused by crop type rotations, fallow practices, and annual climate differences, applying generalized classifiers to years without samples is often problematic, especially for those years with adequate precipitation when rainfed croplands can appear just as green as irrigated ones (Deines et al., 2019; Ozdogan and Gutman, 2008; Xie et al., 2019). As such, methods that can generate training samples on a yearly basis would be optimal for annual irrigation mapping. Xie et al. (2019) proposed a thresholding method to automate the procedure of training sample collection in the western U.S. for a drought year of 2012. The method, however, calibrates crop greenness thresholds based on USDA-NASS irrigation census data, which is available only every five years. Although promising, extending the method to non-census and high precipitation years has thus far remained unresolved. Even more restricting is the dearth of training samples in the eastern U.S., which continues to prevent adequate mapping of the region, despite the area's disproportionate increase in irrigation extent.

Unsurprisingly, there is a strong west-east division in the complexity of mapping irrigation across the U.S. – generally simpler in the west and more challenging in the east (Brown and Pervez, 2014; Ozdogan and Gutman, 2008; Pervez et al., 2014; Xie et al., 2019). Temporally, distinguishing irrigated from rainfed croplands is particularly challenging in high precipitation years and wet regions (Karthikeyan et al., 2020; Pageot et al., 2020). However, multiyear patterns of irrigation deployment and use and their relationship to interannual climate fluctuations may lend novel insights capable of aiding the classification process. For example, those fields equipped with irrigation systems are likely to be cropped and irrigated more frequently than unequipped fields due to the added infrastructure investment and associated costs. These logic-based approaches may hold promise to improve overall mapping quality of annual products, particularly for the locations where the quality of automatically generated training samples is low.

This study proposes a method to produce annual 30 m irrigation extent maps across CONUS from 1997 to 2017, named the Landsat-based Irrigation Dataset (i.e., LANID-US 1997–2017). The work builds upon previous efforts presented in Xie et al. (2019) to overcome several key challenges including its single-year nature, limited training data in the east, inapplicability of the proposed method to non-census years, and limited accuracy assessment. To overcome previous obstacles, we developed an automated training sample generation method in the West that is capable of ameliorating data gaps in non-census years. We paired this with a simplified sample collection strategy in the East to surmount the limiting factor of training and validation availability in other studies to date. In addition, we propose an irrigation practice-based post-classification method based on each annual product. Our developed method was implemented on Google Earth Engine (Gorelick et al., 2017) and is

easily extendable to other mapping years and geographic locations.

2. Methodology

In this study, we mapped irrigation extent across CONUS at 30 m resolution (see Text S1 for detailed study area description, Fig. 1). We defined irrigation as croplands (excluding pasture/grass and non-alfalfa hay) that fully or partially relied on artificial supply of water to offset water stress during the entire growth period of a given year. After describing the datasets and input features used in this research, we detail the methods used for: (1) generating a mask of maximum cropping extent during 1997–2017; (2) preparing a training data pool; (3) designing and conducting the classification; (4) post-classification processing; and finally, (5) evaluating map accuracy (Fig. 2).

2.1. Datasets, processing, and input features

We used multi-source datasets in this study, including satellite imagery as well as environmental and climate variables. We utilized Landsat 5, 7, and 8 images (C1 Level-1 surface reflectance, 16-day revisit frequency) captured between 1997 and 2017, with the annual image availability ranging from 13,000 to 15,000 except for 2012 and the years before 2000 (around 7500). These data have been atmospherically corrected and contain a mask of cloud, shadow, water, and snow (Zhu et al., 2015; Zhu and Woodcock, 2012). In addition, we used 250 m MODIS-derived enhanced vegetation index (EVI) (16-day composites, MOD13Q1 Version 6) and 1 km land surface temperature (LST) (8-day composites, MOD11A2 Version 6) from 2001 to 2017, as these products are useful for classifications in the eastern CONUS (Xie et al., 2019). Compared to Landsat data, MODIS observations provide a more complete profile of intra-annual agricultural cycles despite its coarser spatial resolution (Chen et al., 2018; Didan 2015; Ozdogan and Gutman, 2008). The EVI composites are made using the highest EVI value with low cloud coverage and low view angle from all the MODIS acquisitions within each 16-day period. Each daytime and nighttime LST pixel is the mean of all the corresponding MOD11A1 pixels collected within that 8-day period. Lastly, we used the 1 km resolution daily surface weather and climatological summaries dataset (Daymet V3, Thornton et al., 2016) for climate variables that are closely related to irrigation activities, i.e., precipitation, air temperature, and partial pressure of water vapor (Deines et al., 2019; Deines et al., 2017; Jackson et al., 1981; Ozdogan and Gutman, 2008; Thornton et al., 2016; Xie et al., 2019). The Daymet V3 dataset provides gridded estimates of daily weather parameters for the U.S., derived from selected meteorological station records and various supporting sources (Thornton et al., 2016). Other data sets used in this study consisted of top 25 cm soil water holding capacity extracted from the U.S. SSURGO soil database (NRCS, 2016), U.S. major rivers (Fig. 1, Esri (2010)), and 30 m digital elevation model (Farr et al., 2007). All these data sets were accessible on Google Earth Engine, except for the soil database and major rivers, which were accessed and uploaded manually.

After removing cloud-contaminated and saturated pixels in each image, the maximum, median, and range composites of greenness index (GI), EVI, and normalized difference water index (NDWI) were calculated annually by using all Landsat images available within each year. To eliminate the effect of unidentified cloud and poor-quality pixels, the maximum and minimum values for each pixel were calculated at the 95th and 10th percentiles, respectively. The MODIS-derived products (i.e., EVI and LST) were temporally smoothed and aggregated to annual and late season (May 1 to October 15) sum. Annual maximum EVI was computed from MODIS products as an additional input. In addition to the Landsat and MODIS-based indices, we further utilized several environmental variables in classification (Deines et al., 2019; Deines et al., 2017; Xie et al., 2019; Xu et al., 2019). Specifically, this included annual and late season sum of select climate variables (i.e., precipitation, temperature, partial pressure of water vapor) and some static variables such as elevation and slope, soil water content for the top 25 cm of soil, and distance to major rivers. Furthermore, we adopted

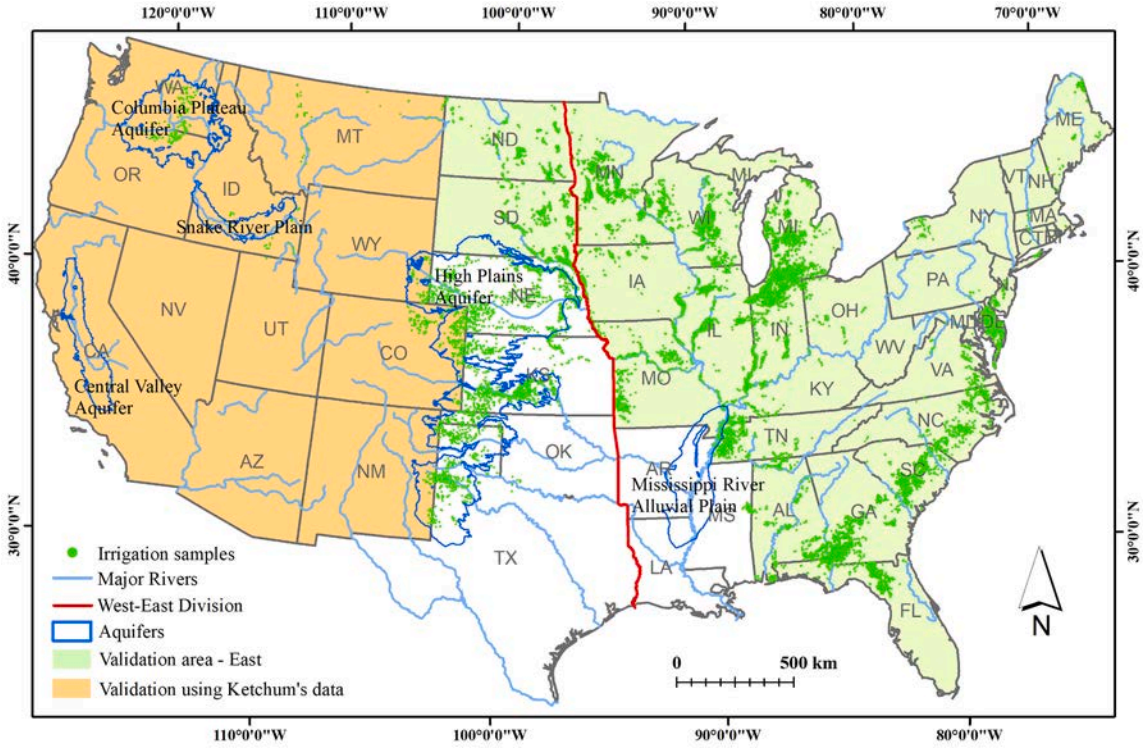

Fig. 1. Study area and distribution of training and validation samples. Irrigation samples in this figure represent the manually collected ones, while automatically generated samples are not shown. The West-East division is based on a climatic transition near the 100th meridian: the more arid West and the less arid East (Seager et al., 2018).

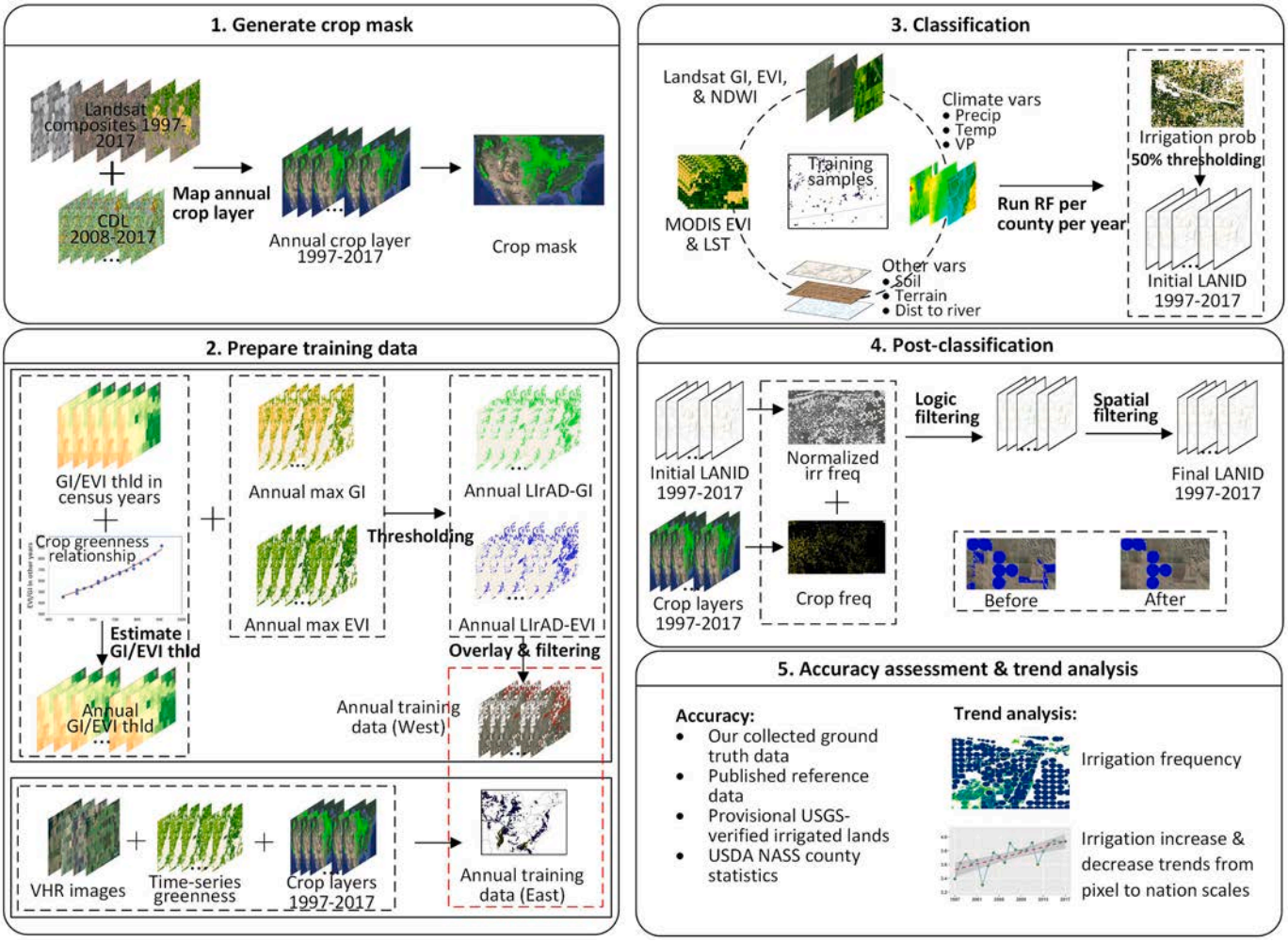

Fig. 2. Methodology overview.

several composite indicators by combining remote sensing indices and environmental variables as they help highlight contrasts between irrigated and rainfed croplands, namely the water adjusted greenness index (WGI) and the aridity normalized greenness index (AGI) (Deines et al., 2017). Altogether, there were 32 input features except for the years before 2001 (i.e., 1997–2000) when MODIS products were not available (Table S1). All input variables were resampled to the 30 m resolution for further use.

2.2. Creating crop masks

Because our automated sample generation method required removal of non-cropped locations, we first created annual nationwide maps of cropland extent by using a generalized classification approach (see Text S2 for details). Although maps of cropland extent were available in selected years from National Land Cover Database (NLCD) and Cropland Data Layers (CDL) at the time of this study, they did not adequately cover the years from 1997 to 2007 (Boryan et al., 2011; Yang et al., 2018). As assessed by CDL, which has cropland accuracies greater than 97% (Lark et al., 2017, 2021), our generated annual maps had a reasonable overall crop vs. non-crop accuracy of 89.7% over the country (Table S2). The resulting maps were subsequently stacked to create the maximum crop extent during 1997–2017 (i.e., crop mask).

2.3. Generating irrigation vs. non-irrigation training data

One of the biggest challenges for accurately mapping irrigated croplands is preparing training data that spatially cover the entire CONUS and are also consistent over time. In this study, we designed two sets of training data—an automated method for the western CONUS and a simplified manual sampling strategy for the eastern CONUS (Fig. 1). Together, these approaches contributed to a training pool that was quantitatively sufficient for yearly classification.

To generate training samples for the western states (including Arkansas, Mississippi, and Louisiana) in USDA-NASS census years (i.e., 1997, 2002, 2007, 2012, and 2017), we adopted the thresholding method proposed by Xie et al. (2019), which assumed that irrigated croplands appear greener than those that are rainfed (Pervez and Brown, 2010; Pervez et al., 2014). The method first generated two intermediate Landsat-based Irrigation Agriculture Dataset maps (LIrAD-GI and LIrAD-EVI) by using a county-constrained thresholding method, where agricultural pixels (identified from maximum crop mask) with yearly maximum GI (Gitelson, 2005; Ozdogan and Gutman, 2008) or EVI (Liu and Huete, 1995) greater than the optimal threshold were labeled as potentially irrigated croplands. Irrigation sample candidates were then selected as those pixels classified as potentially irrigated in both LIrAD-GI and LIrAD-EVI, while non-irrigated samples were those classified as non-potentially irrigation in both maps. To calibrate county-specific optimal thresholds of GI and EVI, we iteratively segmented GI (or EVI) by using a series of thresholds until LIrAD-GI (LIrAD-EVI) resulted in the same irrigated area as reported by USDA-NASS (refer to Xie et al. (2019) for details).

While this sample generation method provided sufficient samples for census years, it is inapplicable to non-census years when USDA-reported

irrigated area is not available. Simply applying the optimal thresholds of census years to non-census years is also problematic, however, because thresholds are not temporally comparable due to differences in crop type and growth trajectories, climate, and availability of clear Landsat observations across years. For example, in our testing, we found that applying the threshold of 2007 to other census years would lead to substantial mis-estimation of LIrAD and thus would introduce considerable errors to the sample datasets (Fig. 3). Moreover, irrigation decisions are temporally dynamic, so it is not practical to use training samples of a single year as a "stable class" for other years.

To overcome these challenges, we developed a normalization method whereby the relationships of crop greenness between years was used to estimate optimal GI and EVI thresholds for non-census years from those of census years. We modeled the relationship of cropland greenness (*vi,* represented by yearly maximum GI and EVI) between a non-census (*NCYr*) and census year (*cYr*) for the *i*th county as:

$$vi_{i,NCYr} = a_{i,cYr} + b_{i,cYr} * vi_{i,cYr} + c_{i,cYr} * vi_{i,cYr}{}^2 \quad (1)$$

where *a*, *b*, and *c* are year- and county-specific coefficients. To obtain these regression coefficients in an efficient manner, we used county-level percentile values of GI (EVI) (from 5 to 95 percentile with an increment of 5 points) instead of all crop pixels. In addition to saving computation time, this percentile-level regression further served to reduce the impacts of inter-annual differences in image quality, i.e., cloud contamination and Landsat ETM+ scanline corrector-off effect.

Next, the optimal GI (EVI) threshold of a county *i* in a non-census year was predicted as:

$$estThld^{vi}_{i,cYr} = a^{vi}_{i,cYr} + b^{vi}_{i,cYr} * optThld^{vi}_{i,cYr} + c^{vi}_{i,cYr} * optThld^{vi}_{i,cYr} 2 \quad (2)$$

where $optThld_{i,\ cYrvi}$refers to the optimal GI (EVI) threshold in a census year *cYr*. Given that multiple thresholds can be calculated using different census years as references, the final threshold of a non-census year ($optThld_{i,\ NCYrvi}$) was estimated as the median value of all five estimated thresholds:

$$optThld^{vi}_{i,NCYr} = median\left(estThld^{vi}_{i,1997}, estThld^{vi}_{i,2002}, estThld^{vi}_{i,2007}, estThld^{vi}_{i,2012}, estThld^{vi}_{i,2017}\right) \quad (3)$$

The estimated thresholds for non-census years were then applied in the same manner as they were for census years to create potential training samples for each year. To further improve quality of training data, the annual potential samples were stacked, and those pixels identified as irrigation candidates for less than three years were deleted. This step is especially important for the years with exceptionally high precipitation when rainfed cropland can be occasionally identified as an irrigated candidate. Finally, a 5 × 5 filter was applied annually to remove mixed pixels along field boundaries. The result was a generated pool of potential training samples that were evenly distributed across the western CONUS on a yearly basis.

For the eastern states, manual collection of training samples was necessary due to the lack of ground truth data and the inefficiency of automatic sample collection methods in the region (Xie et al., 2019). To enable efficient sample collection of irrigated croplands, we focused on selecting center pivot irrigated fields due both to the dominance of this irrigation type and its clear appearance of irrigation infrastructure and indicators on very high-resolution (VHR) images. Specifically, we collected center pivot fields via visual interpretation of the most recent VHR imagery available on Google Earth, based on clear observation of water well heads, irrigation pipe spans, center pivot towers, and/or circular field patterns from the tracks of sprinkler wheels. In total, ~30,000 potential training locations were manually selected across the eastern CONUS. The presence of a center pivot system for 1997–2017 was then recorded through interpreting multi-temporal VHR images on Google Earth and time-series Landsat observations on Google Earth Engine. From the potential training locations, the *i*th selected location in the *yr*th year was used as an irrigation sample if it was cultivated according to our annual cropland extent layers or its greenness was greater than the optimal threshold of the county where it was located:

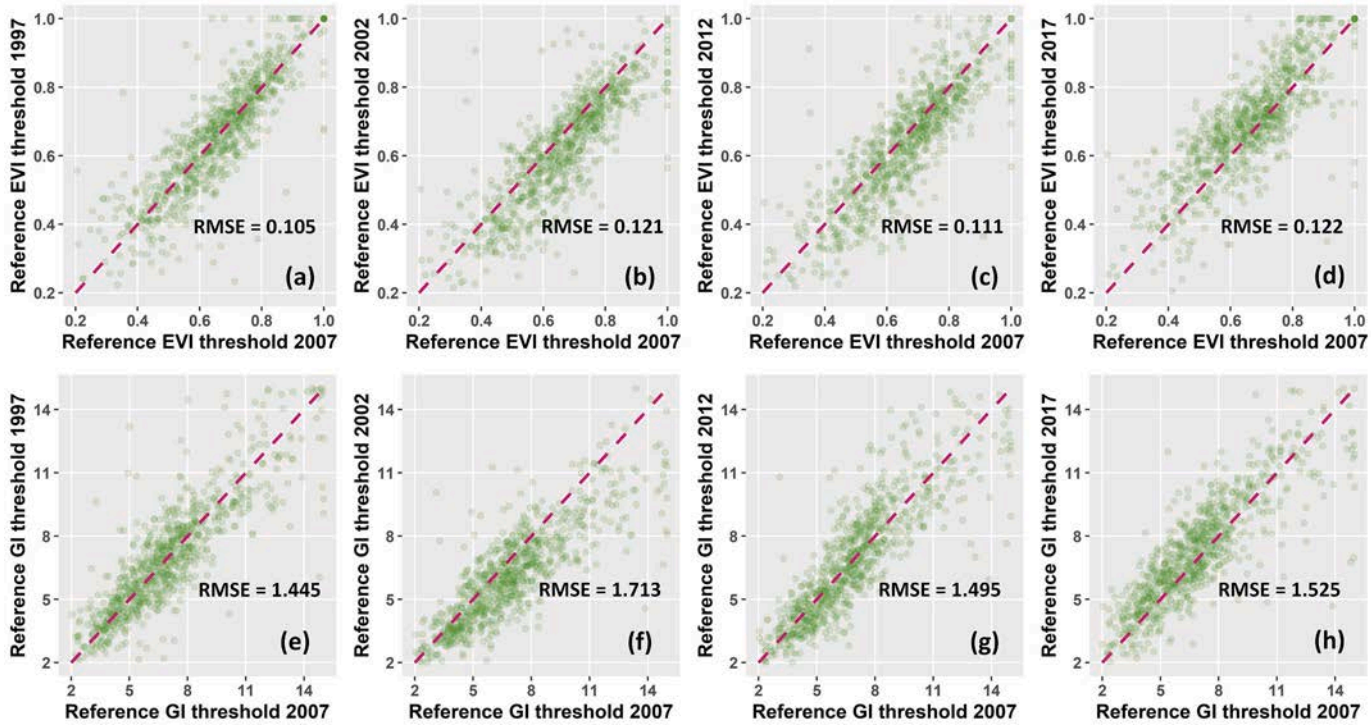

Fig. 3. Demonstration of temporal inconsistency of EVI and GI thresholds. Note that temporal inconsistency of thresholds was found between each pair of census years.

$$training_{i,yr} = \begin{cases} true \quad crop_{i,yr} = true \quad or \quad GI_{i,yr} > \alpha * optThld_{cty,yr}^{GI} \\ false \quad others \end{cases} \quad where\ yr \in [startYr_i endYr_i] \tag{4}$$

where $startYr_i$ and $endYr_i$ refer to the starting and ending year of the presence of an irrigation system for the ith collected location; $optThld_{cty,yr}^{GI}$ represents the optimal GI threshold for the country in year yr, which was determined using the same approach as in the western CONUS. The second term (greenness greater than threshold) was incorporated to reduce the impact of possible inaccurate classification in annual crop layers. The adjustment factor α (set as 0.7 in this study) was applied to possibly include some irrigated samples that can be less green than rainfed fields in humid areas (Xie et al., 2019; Xu et al., 2019).

To collect non-irrigated training data for the east, we randomly extracted 30,000 non-irrigation samples based on LANID 2012 (Xie et al., 2019) and used several pieces of visual evidence to confirm their non-irrigated status throughout time, including the absence of irrigation infrastructure as verified by multi-temporal VHR images (when available) and field greenness and texture contrast with confirmed irrigated fields on Landsat time series. If any evidence of irrigation was observed in either multi-temporal VHR or Landsat-derived images in any year, the sample point was discarded and replaced with a nearby location that could be confirmed as non-irrigated through time.

We also included additional manually collected training locations of permanent irrigation and non-irrigation locations for some semi-arid western states where automated samples can be unreliable (Fig. 1, e.g., North Dakota, South Dakota, eastern Nebraska, Washington, Idaho, and Montana). Here, we recorded locations showing clear irrigated and non-irrigated use through the whole study period and treated them as "stable". The stable non-irrigated locations were used annually, whereas the stable irrigated samples were overlaid with annual crop layers to determine whether they could be used as training in a target year based on Eq. (4). Approximately 80% of manually collected samples in the east and semi-arid western states were used for classifier calibration and the remaining were reserved for accuracy assessment, except for Washington, Idaho, and Montana where all samples were used for training.

2.4. Classification design

The classification was conducted annually per county. The reasons for this are twofold. First, locally adaptive classifiers have higher accuracy than generalized classifiers that apply to broad areas (Johnson, 2019; Massey et al., 2017). This is especially true for irrigation mapping due to spatial variations of crop types and irrigation signals particularly in humid areas. Second, while studies have demonstrated the possibility of applying temporally generalized classifiers through time in arid and semi-arid areas (Deines et al., 2019; Ketchum et al., 2020), we assumed annual classifier training would result in better classifications because feature extension of croplands can be confounded by rotations of crop types over time and differences in crop growth patterns and signals across years.

For classification, we used the random forest classifier due to its reasonable accuracy and ability to handle high-dimensional and multicollinear variables (Belgiu and Drăguţ, 2016; Breiman, 2001; Pal, 2005; Rodriguez-Galiano et al., 2012). For a given year (e.g., 2010), four hundred samples (200 for each class) were randomly extracted from the training pool for those counties where sufficient samples were available, whereas all points were used if the total sample of a class was less than 200 (Fig. S1). These samples and associated input features were used to train a year- and county-specific classifier, which was then applied to generate per-pixel irrigation probability of the broader county area (i.e., county and its 3 km buffer). For the counties with less than 10 samples for any class, a state-wide classifier was applied, which was trained using 4000 samples (2000 for each class, or all samples if the total sample was smaller than 2000) extracted from the training pool. The county-level probability layers were then mosaicked to create a nationwide map (average value was calculated for overlapping areas). The provisional nationwide irrigated area of a year was then identified as pixels with irrigation probability greater than 50%.

2.5. Post-classification

Because our full 1997–2017 maximum crop extent mask included pixels that were non-cropland in some years, areas such as wetlands, forests, and irrigated urban turfgrass were liable to be confused with irrigated croplands and misidentified as such by the classifier. To address this, we used the NLCD broad landcover maps to remove these areas from our provisional irrigation maps. At the time of this study, consistent NLCD maps were available for the years 2001, 2003, 2006, 2008, 2011, 2013, and 2016, which were used to filter irrigation maps between 1997 and 2001, 2002–2004, 2005–2006, 2007–2009, 2010–2012, 2013–2014, and 2015–2017, respectively.

We subsequently applied a logic filter to remove possible false classification caused by occasional overlap in spectral signatures of irrigated and rainfed croplands in high precipitation years. As demonstrated in Fig. 4b–c, we assumed that fields equipped with irrigation infrastructure tend to be both cropped and irrigated more frequently than not, given the requirement and costs of infrastructure. Thus, we proposed the normalized irrigation frequency as an indicator to remove potential overestimation of the irrigated class.

$$normIrrFreq = (irrFreq) / (irrYr_{last} - irrYr_{first}) \tag{5}$$

where $irrFreq$ is the number of years a pixel is mapped as irrigation in the provisional LANID time series, and $irrYr_{last}$ and $irrYr_{first}$ are the last and first year a pixel is detected as irrigation. The values of $normIrrFreq$ range from 0 to 1, with higher values representing more frequent irrigation within the timespan that irrigation was provisionally detected. We then relabeled as non-irrigated any pixels with normalized irrigation frequency less than 0.5. To avoid over-filtering, however, we did not implement this post-classification refinement for any fields that had a normalized cropping frequency above 0.5, defined as:

$$normCropFreq = cropFreq / Yr_{total} \tag{6}$$

where cropFreq is the number of years a pixel is identified as crop in annual crop layers and Yr_{total} is the length of study period (i.e., 21). The combination of these two rules removed irrigated pixels in the provisional LANID that were both cropped and irrigated infrequently. We found this logic rule to be especially valuable in areas with frequent fallow (Fig. 4a, d).

Finally, a spatial filter was applied to each annual map to further remove classification noise based on the assumption that small fields are unlikely to be irrigated given the requirement of irrigation infrastructure. Therefore, we relabeled isolated irrigation clusters smaller than 23 Landsat pixels as non-irrigated and filled in small gaps (less than 2 ha) classified as non-irrigated within broader irrigated fields. The resulting, final LANID product consisted of 21 nationwide maps of binary irrigation status.

2.6. Accuracy assessment

Accuracy assessment of irrigation maps across a broad area can be challenging due to limited availability of existing ground truth samples and the high resource costs of collecting new, reliable data. Here, we integrated several reference datasets to provide a comprehensive nationwide assessment of annual irrigation maps in the U.S. Reference data included visually collected samples in this study, published data from previous studies, provisional USGS-verified irrigated lands, and USDA-NASS reported county statistics. Our visually collected samples were withheld and selected from the training data described in Section 2.3; Published reference data included those from Ketchum et al. (2020)

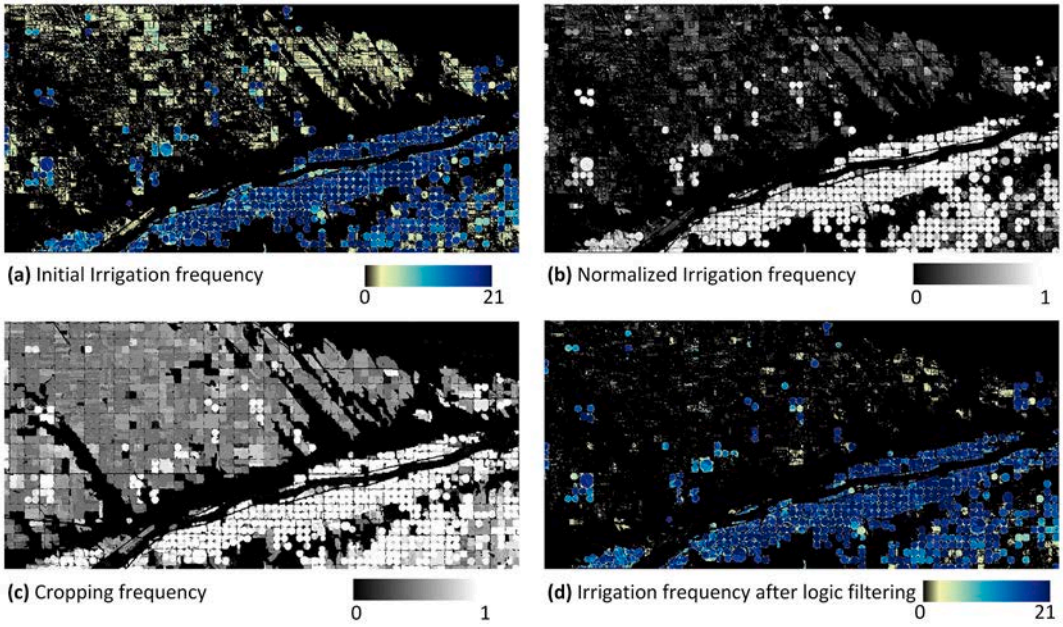

Fig. 4. Demonstration of logic filtering in the boundary area of Deuel, Keith, and Garden County of Nebraska. Irrigation and cropping frequency refer to how often a field has been irrigated and cultivated during the study period, respectively.

and Deines et al. (2017) for the eleven western states and the northern High Plains Aquifer, respectively (Fig. 1); A provisional version of the USGS-verified field polygon data were compiled from various state and federal agencies (Brandt et al., 2021); and USDA-NASS statistics were derived from the Census of Agriculture. First, county- and state-level irrigation areas aggregated from LANID were compared to USDA-NASS data for the census years (i.e., 1997, 2002, 2007, 2012, and 2017). We then assessed LANID's pixel-level accuracy using our test samples from this study, Ketchum et al. (2020), and Deines et al. (2017). We finally compared LANID with USGS field reference data. Since USGS-verified data focus on the western CONUS and were only available for recent years, they were used to assess selected recent maps (Table 1). One thousand test samples (500 for each class except for Montana) were randomly stratified from irrigated and non-irrigated lands for each validation year. Table 1 lists the number of irrigated and non-irrigated test samples for each evaluation area.

Table 1
The number and source of test samples for each validation year. Note spatial coverage of provisional USGS-verified irrigated lands within each state varied across time.

States	Validation years	Annual number of samples		Source
		Irrigated	Non-irrigated	
Nebraska	1997–2017	2400	4858	This Study (visually collected)
Kansas	1997–2017	1736	1638	
Texas	1997–2017	688	1292	
New Mexico	1997–2017	146	241	
Colorado	1997–2017	600	625	
East	1997–2017	5000	5000	This Study (visually collected)
Eleven western states	1997–2017	606	1072	Ketchum et al. (2020)
NHPA[a]	2002	229	790	Deines et al. (2017)
	2015	210	1013	
Arizona	2013, 2014, 2016	500	500	USGS (Brandt et al., 2021)
Colorado	2016	500	500	
Idaho	2011, 2015	500	500	
Montana	2013	200	100	
Utah	2009	500	500	
Washington	2015–2017	500	500	

[a] The Northern High Plains Aquifer.

3. Results

3.1. Thresholds estimation, quality of training samples, and variable importance

We found strong year-to-year discrepancies among EVI and GI thresholds when attempting to apply thresholds calibrated in a census year to other years (Fig. 3). In contrast, our proposed normalization method provided improved prediction of EVI and GI thresholds, with lower RMSE values between the reference and estimated threshold values as compared with those without normalization (Fig. 5). The normalized estimations yielded RMSE of ~0.08 and ~1.1 for EVI and GI, respectively, compared with ~0.1 and ~1.4 for the non-normalized, surrogate-year approach. Threshold estimation of a non-census year should be even more accurate and robust since all census years can be used as reference instead of withholding one for comparison, as done for analysis of Fig. 5. The good agreements between estimated and reference thresholds indicate that those for non-census years can be accurately predicted, thereby providing a sufficient approach for training sample generation for these years.

The overall accuracy of the derived annual training samples was upwards of 90% across the western CONUS (Table 2) using the normalized threshold approach to automated sample collection. We further found that the quality of the automatically extracted samples was better in drought years (e.g., 2012) compared to non-drought years (e.g., 2009 and 2015), given the greater contrast in greenness between irrigated and rainfed croplands. We also evaluated one of our training sample selection criteria for the eastern CONUS by testing the assumption that a field equipped with a center pivot irrigation system would likely to be irrigated if it were cultivated (Eq. (4)). We randomly selected fifty center pivot fields across the eastern states and checked for irrigation application on growing season VHR images available on Google Earth. We found that 149 out of 155 growing season observations (3.1 clear observations per field during 1997–2017) showed strong indicators of irrigation, such as the appearance of concentric circular tacks from wheel movement. This high rate of irrigation application for fields

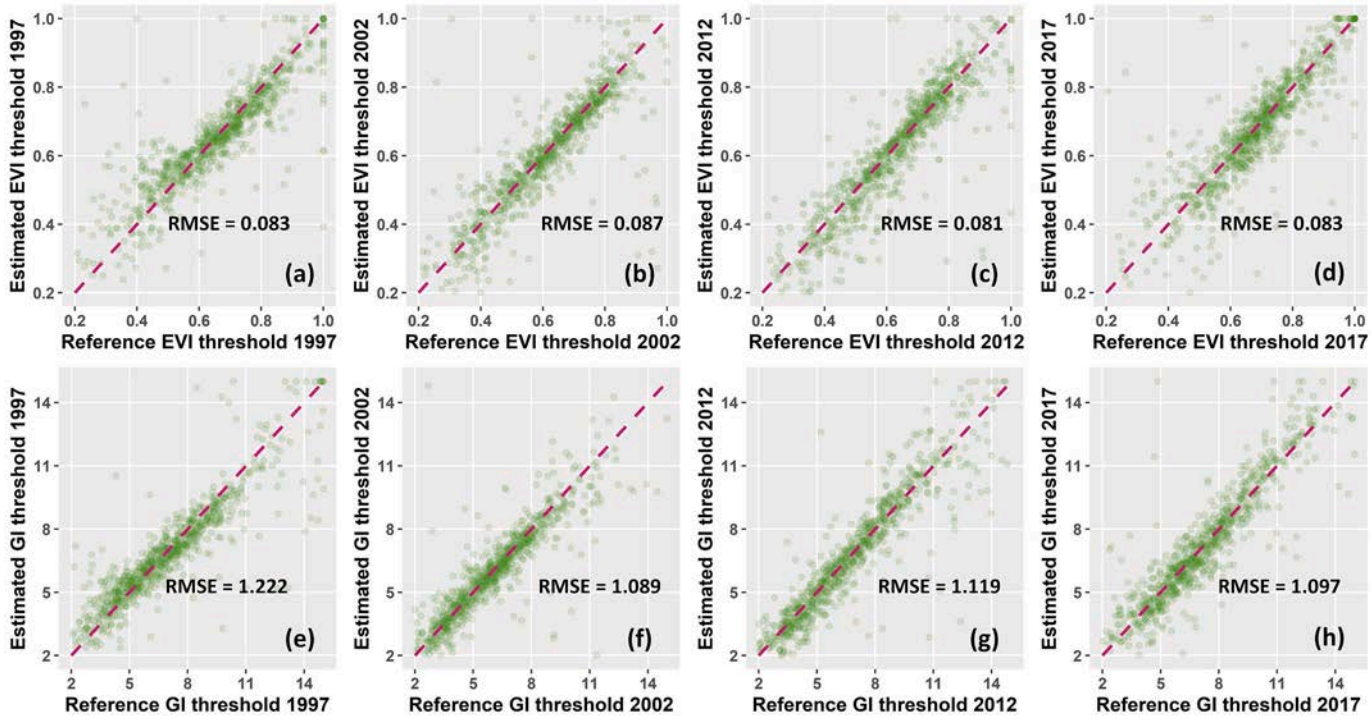

Fig. 5. Demonstration of threshold estimation using the proposed normalization approach. For each subplot (a-h), the estimated threshold was predicted based on data from the other four census years (e.g., 1997 was estimated using 2002, 2007, 2012, and 2017) and compared to the actual (reference) threshold for the target year. General agreement between the estimated and reference data, especially compared to the non-normalized surrogate-year approach (Fig. 3), support use of the normalization routine.

Table 2
Accuracy of extracted training samples in the western CONUS for representative years.

		Ground truth					
		2009		2012		2015	
		Irrigated	Non-irr	Irrigated	Non-irr	Irrigated	Non-irr
Training sample	Irrigated	46	6	52	7	43	5
	Non-irr	1	62	0	72	4	47
OA		93.9%		94.7%		90.9%	
Kappa		0.88		0.89		0.82	

equipped with center pivot irrigation system further supports our approach to training sample generation in the eastern CONUS, with generated samples covering a wide range of greenness and crop types (Fig. S2).

Generally, Landsat-based variables (e.g., GI, EVI, and NDWI) were likely to rank among the most important variables at the county scale for different precipitation levels (Fig. 6), although their contributions varied slightly across county (Fig. S3). Climate variables tended to play low to moderate roles with high spatial variation, possibly due to similar climates within a county. Environmental and climate variables gained increasing importance at the state level (e.g., temperature, precipitation, available water content, and elevation), especially for the humid east (e.g., Minnesota in Fig. S4). It is also evident that the importance of a feature can change over time for a specific county, with some showing large temporal variations such as MODIS-derived EVI variables (e.g., late season MODIS EVI for Adams County, WA and the maximum MODIS EVI for Finney County, KS, Wadena County, MN, and Mitchell County, GA).

3.2. Qualitative assessment of LANID

Our classification generated 21 annual nationwide irrigation maps for the years 1997 to 2017. For visual comparison, we present the results of the LANID in Fig. 7, along with similar views from IrrMapper (Ketchum et al., 2020), Annual Irrigation Maps – High Plains Aquifer (AIM-HPA) (Deines et al., 2019), and MIrAD-US (Pervez and Brown, 2010). From initial inspection, the 2017 LANID map captures key irrigation hotspots across the U.S. (Fig. 7) and matches the broad irrigation patterns reported in other regional or coarser-resolution studies (Brown and Pervez, 2014; Deines et al., 2019; Ketchum et al., 2020; Ozdogan and Gutman, 2008; Pervez and Brown, 2010).

At local scales, however, the LANID data provides improved depiction of irrigated field shapes and boundaries than other nationwide maps while showing similar patterns and details as the regional IrrMapper and AIM-HPA. For example, in the central Washington and High Plains Aquifer, the 30 m LANID 2017 map effectively captures field-level details of well-established irrigation clusters as well as scattered individual fields outside intensively irrigated hotspots (Fig. 7a, b). The 250 m MIrAD-US generally captures the same irrigated areas but with fewer details. Outside of the arid western regions, the differences between maps are amplified and extend beyond level of detail. In central Wisconsin, for example, our LANID map appears to substantially outperform MIrAD, which tends to identify the greenest pixels. (Fig. 7c). Despite this notable improvement in both resolution and accuracy, however, the

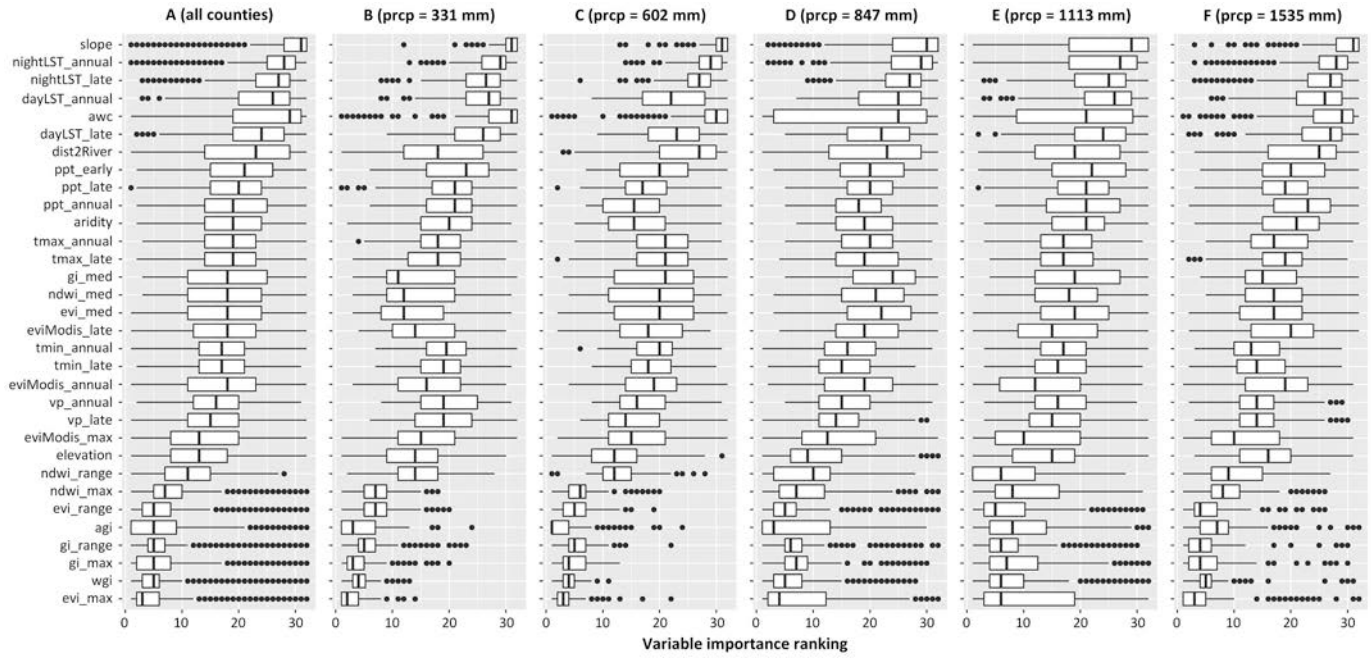

Fig. 6. County-level variable importance ranking for the year 2017. A: overall ranking, B-F: grouped by precipitation levels. Around 1500 counties were evaluated. The numbers in parentheses are the average total precipitation of each group (n ~ = 300). The most important variables are listed towards the bottom of the chart, with mean county ranks closest to 1. EVI_max was the most important variable overall.

LANID map does not distinguish individual fields in the region as clearly as in the High Plains Aquifer due to weaker contrast between irrigated and rainfed fields and more frequent use of corner irrigation in central Wisconsin than in the High Plains Aquifer.

By stacking the LANID annual irrigation maps, we further identified the frequency of cropland irrigation (i.e., number of years a field was detected as "irrigated") across CONUS (Fig. 8). As expected, heavily irrigated areas (dark blue in Fig. 8) are particularly prevalent in several aquifers including the High Plains Aquifer, the Mississippi Alluvial Plain, the Central Valley Aquifer, the Snake River Plain, and the Columbia Plateau Aquifer. In addition, there are select areas with considerable and frequent irrigation in the eastern CONUS, including the Upper Midwest (e.g., Minnesota, Wisconsin, Michigan, Illinois, and Indiana) and the East Coast (e.g., Georgia, Florida, South Carolina, and Delaware).

Field-level inspection of irrigation frequency further reveals insights into the dynamics of irrigation activity and the intensity of annual utilization across space. For instance, light blue and green fields in Fig. 8d-f are not as intensively irrigated as those under dark blue, implying that they are either newly irrigated, recently abandoned, or fields equipped with irrigation system but under frequent fallow or non-use. Stacking all 21 annual maps can also aid identification of irrigation infrastructure types - such as center pivot sprinkler systems, as indicated by clear circular field shapes even in the humid eastern areas (Fig. 8g-i). This capability is particularly pronounced when compared to single year maps, in which mapping errors may obfuscate easy visual identification.

3.3. Quantitative assessment of LANID

The LANID maps provide an accurate county- and state-level areal estimate of irrigated croplands as compared to the USDA-NASS census data, with R^2 over 0.95 for all years of overlap (Fig. 9). Overall, LANID slightly underestimates county-level irrigation area relative to USDA-NASS data (except for the year 2007) and the results are not sensitive to outlier counties with extremely large irrigation area (Fig. S5). Because LANID excludes irrigated grass and hay, the differences are particularly notable in pasture/grass and hay dominated counties, such as in Jackson County, Colorado, and Albany County, Wyoming. We do not find overall over- or underestimation bias at the state level, although LANID does under-represent irrigated area in some western states with considerable amounts of irrigated pasture and hay. Because these states possess relatively smaller amounts of irrigated area, however, they did not impact the overall state-level estimation.

On a per-pixel basis, the overall accuracy (94–99%) indicates that the LANID maps provide a high-quality characterization of irrigation distribution as evaluated by our visually collected ground truth data (Table 3). Although our maps show lower overall accuracy in the eastern states compared with the western, we expect the reduction in map accuracy there to be less than that of other maps covering the entire CONUS. For example, our new 2012 map in this study shows a remarkable increase of accuracy compared to its predecessor reported in Xie et al. (2019), with OA of 94.4% vs. 74.4%. Importantly, the current LANID shows good quality in the U.S. Corn Belt (e.g., Illinois, Iowa, Indiana, Michigan, Wisconsin, and Minnesota), with OA over 90% (Table S3). Furthermore, our maps have an OA above 94% as assessed by published reference data, demonstrating that the LANID maps provide comparable performance with IrrMapper and AIM-HPA in the western U.S. while also providing similarly high-quality identification across the eastern half of the country. It is also evident that the accuracy of our maps is consistent over time (Tables S4–5).

Producer's and user's accuracies present a more nuanced view of how the maps depict irrigated croplands across regions and through time. Similar to overall accuracy, the producer's accuracy (1 – omission error [OE]) for our LANID maps show reasonable performance with mean OE values of 2.0 to 17.1% for the irrigation class across testing areas, suggesting that our approach did not substantially miss (or underestimate) irrigation extent on the ground. On the other side, commission errors (CE) ranged from 0.6 to only 11.3%, indicating very high user's accuracies and that our map does not over-classify irrigation areas or mistakenly label non-irrigated areas as irrigated.

The LANID maps show less consistent results compared with provisional USGS-verified irrigated lands, having OA ranging from 76.7 to 93.6% (Table 3). Despite fair consistencies, a noticeable amount of USGS-verified irrigated lands are not detected in LANID (reflected by high omission error of 10.4 to 39%) and our mapped non-irrigated

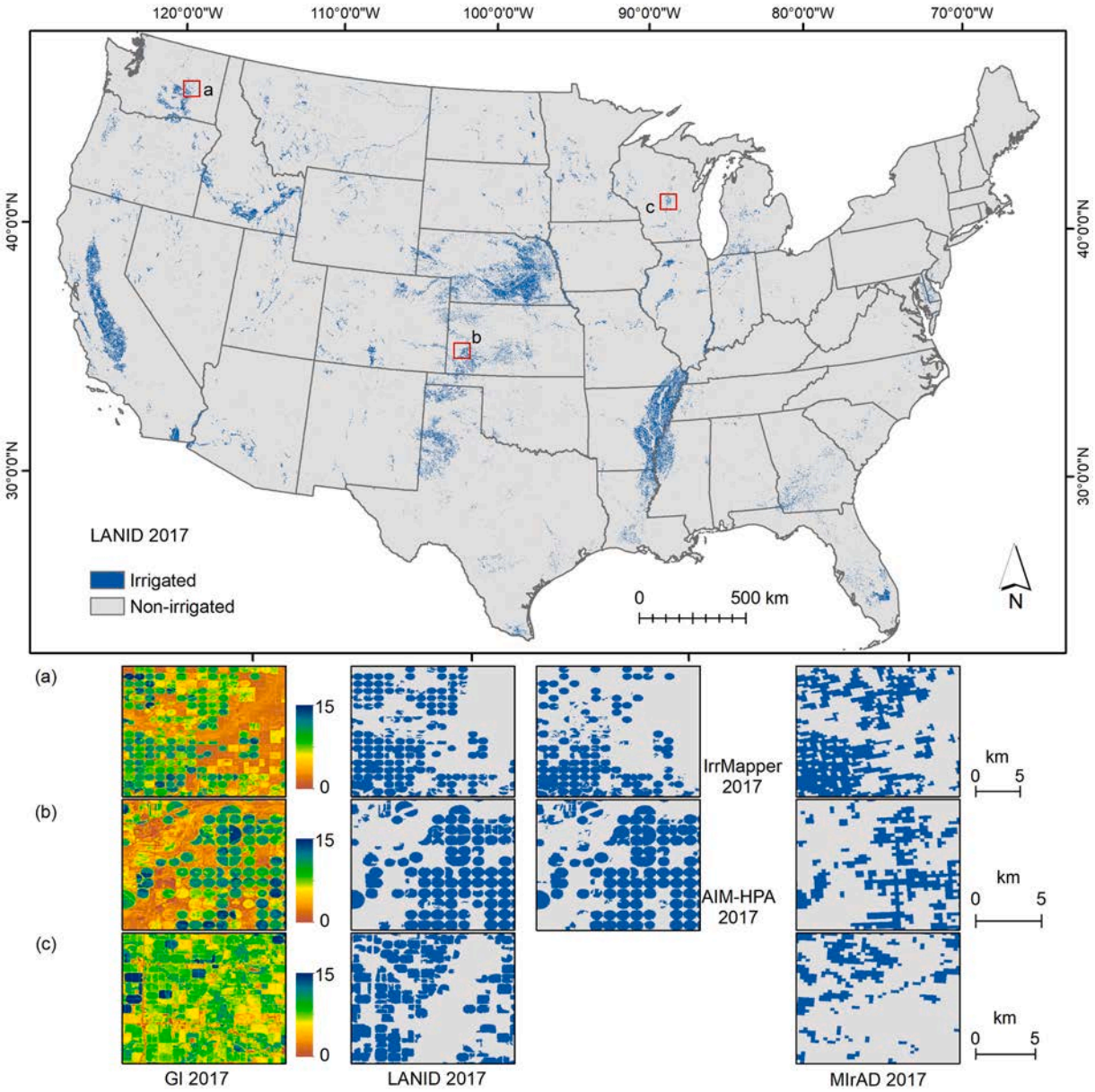

Fig. 7. LANID 2017 (the overview map) and its comparisons with other maps at central Washington (a), southwestern Kansas (b), and central Wisconsin (c). GI: Greenness Index, AIM-HPA: Annual Irrigation Maps – High Plains Aquifer (Deines et al., 2019), MIrAD: MODIS-based Irrigation Agriculture Dataset (Brown and Pervez, 2014; Pervez and Brown, 2010). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

extent mistakenly covers 9.6 to 29.5% of USGS-verified irrigated lands. These disagreements can be attributed to several differences. First, the provisional USGS-verified irrigated lands were based on field boundaries provided by state and federal agencies that were digitized in years different from the time of irrigation verification; the delineated polygons are also more indicative of water use rights rather than of actual irrigation application. As such, the product may not accurately depict real-time irrigation status at sub-field to field levels (Fig. 10). Second, the USGS-verified dataset is not comprehensive – e.g., while those fields mapped as irrigated based on water rights permits are those most likely to be irrigated, not all irrigated lands are captured by the dataset. Third, unlike the USGS data, pasture and non-alfalfa hay are not included in our LANID maps, resulting in considerable inconsistency in areas dominated by these grass-like land covers such as the validation area of northeastern Utah. Lastly, the LANID maps could be inaccurate for some high precipitation years when rainfed croplands have high probability to be incorrectly classified.

3.4. Irrigation trends of CONUS

The annual LANID maps reveal diverse local to nationwide irrigation trends across time. A nationwide analysis of total irrigated area indicates that CONUS has seen a substantial expansion of irrigated croplands during the past two decades, increasing by 1.7 million hafrom ~22.3 in 1997 to ~24 million hain 2016 (Fig. 11a). The average annual increment of countrywide irrigation was approximately 82,000 ha., 84.5% of which occurred in the East (~69,400 ha per year, Fig. 11b). Although Nebraska had the largest irrigation gain in the West (over 20,000 ha per year, Fig. 11d), the entire western CONUS did not show a significant change (Fig. 11c).

Notable increases of irrigated area were observed especially in the Upper Midwest instates such as Minnesota, Michigan, Wisconsin, Illinois, Indiana(Fig. 12a). Despite their already heavily irrigated status, several states in the Mississippi Alluvial Plain region also experienced additional expansion of irrigation. California, Colorado, and Utah

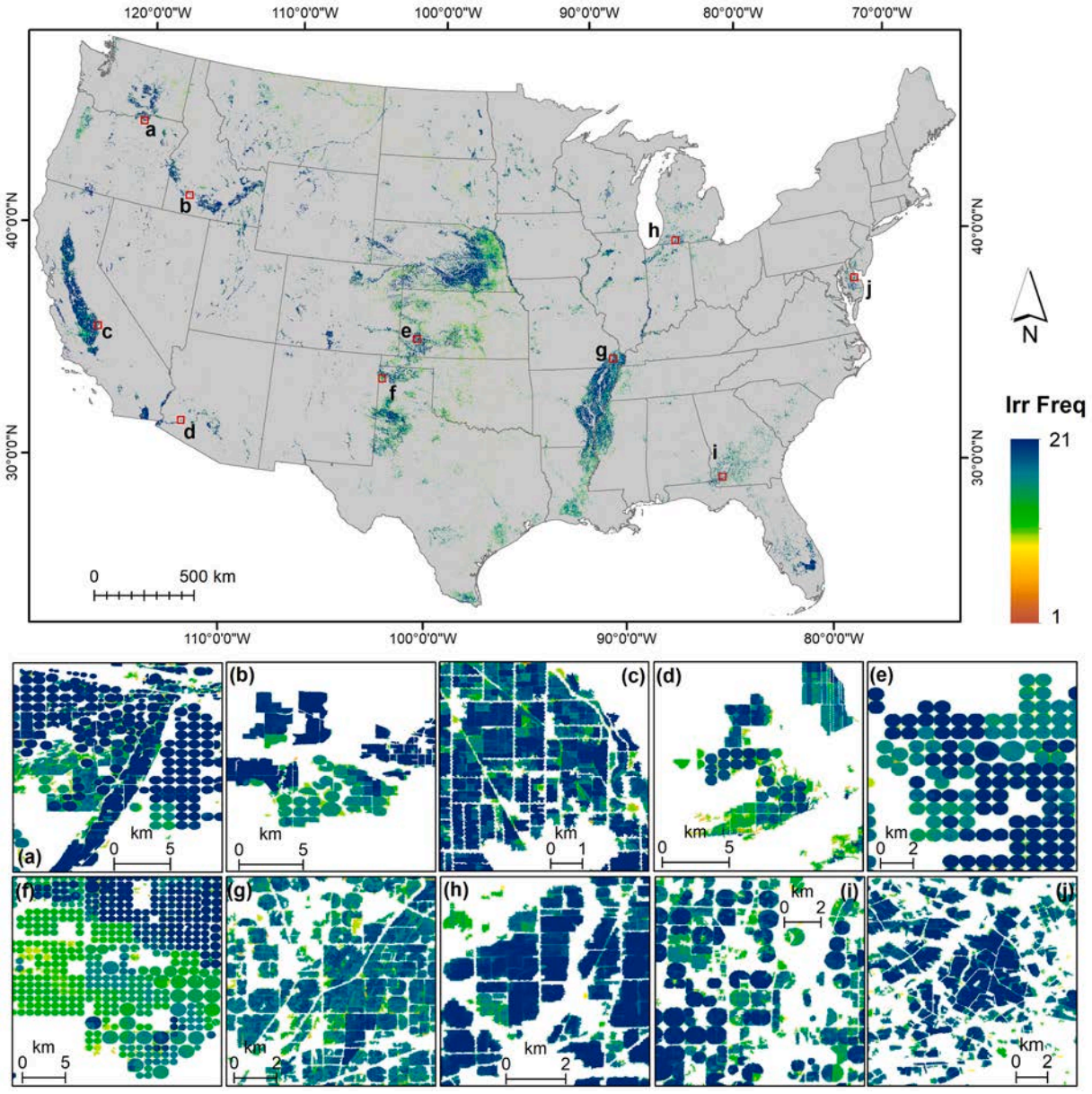

Fig. 8. Nationwide and local views of irrigation frequency, i.e., the number of years irrigated during 1997–2017.

experienced statistically significant contractions of irrigated area, with rates of decline ranging from 300 to 7000 ha per year. Further assessing LANID at the finer, county level reveals more concentrated changes including large expansions of irrigation along the Mississippi River and in eastern Nebraska, central Minnesota and Wisconsin, southern Michigan, northern Illinois and Indiana, Delaware, and central South Carolina (Fig. 12b). Pixel-level examination further indicates that irrigation increases in these areas were frequently along rivers or in locations with sufficient groundwater recharge, suggesting that easy access to ample water resources has largely accommodated irritation expansion in the eastern CONUS (Fig. 13). In addition, we found areas with sizeable expansion in northern Texas (e.g., Hartley County), even though the central and southern High Plains Aquifer has been experiencing groundwater depletion over recent decades (Brown and Pervez, 2014; Scanlon et al., 2012). Only a few other western counties showed large increases, such as Benton County, Washington, and Harney County, Oregon. Hotspots of irrigation decrease were mainly observed in the central and southern High Plains Aquifer (e.g., eastern Colorado, western Kansas, and northern Texas), southern California and Arizona, and southeastern Florida. Sub-grid investigation in Fig. 13 also shows the ability of our LANID maps to capture field-level irrigation change over time, such as abandonment of irrigated fields in central Idaho and irrigation expansion in northern Texas and central Minnesota.

4. Discussion

In this study, we developed nationwide 30 m resolution annual maps of irrigated extent for the conterminous U.S. for all years between 1997 and 2017, which represent the highest spatial and temporal resolution time series data available for CONUS to date. This new generation of irrigation products accurately captures – for the first time – U.S. irrigation dynamics at operationally relevant scales including field-level spatial detail and year-to-year variability of use. These advances in utility may in turn facilitate several new threads of inquiry where field-level, crop-specific, annual data would prove useful, such as monitoring of water use and allocation, exploring socioeconomic and environmental drivers of irrigation use and change, and assessing the effects of agricultural irrigation at climate-, crop-, and soil-relevant scales.

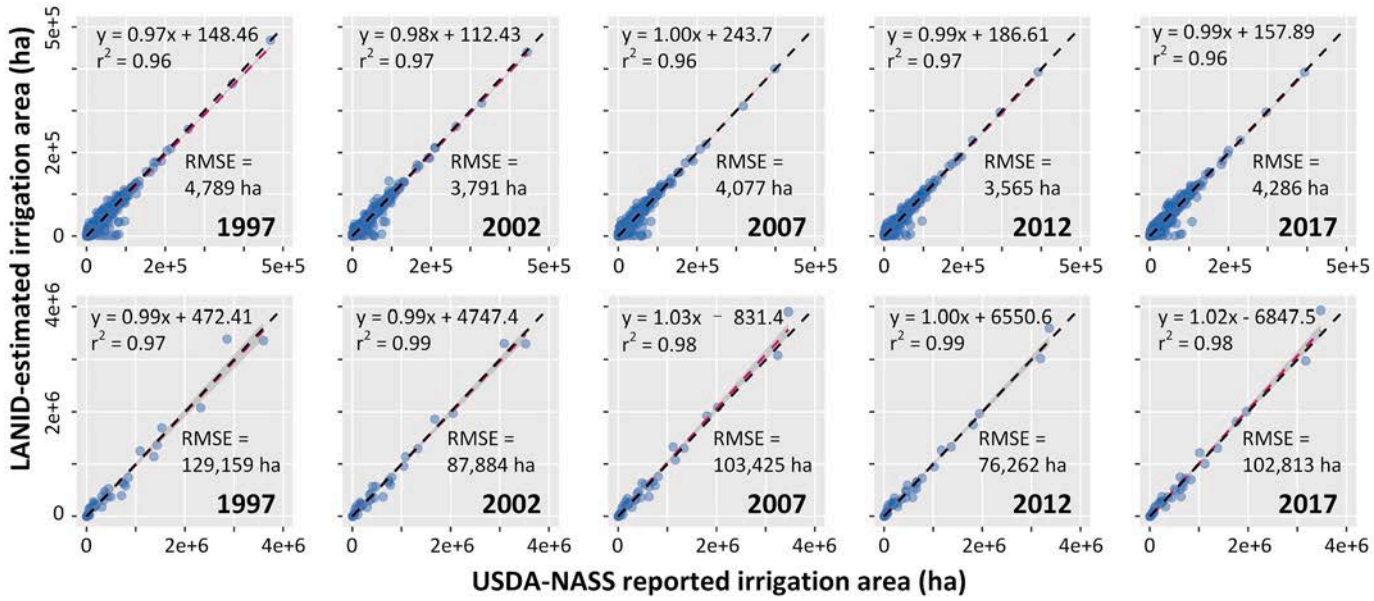

Fig. 9. County- (top row) and state-level (bottom row) comparisons between LANID-estimated and USDA-NASS reported irrigation area for five census years. Note all linear regressions are statistically significant with p value <0.001.

Table 3
Confusion table of pixel-wise accuracy assessment. Accuracy values are averaged if accuracy assessment was conducted in multiple years with parenthetical numbers representing the standard deviation.

Area	Years	Kappa	OA	OE (irrigated)	OE (non-irrigated)	CE (irrigated)	CE (non-irrigated)	Sample sources
Nebraska	1997–2017	0.91 (0.02)	96.2 (1.0)	6.3 (1.8)	2.5 (0.6)	5.2 (1.3)	3.1 (0.9)	This study
Kansas	1997–2017	0.91 (0.02)	95.3 (1.2)	8.0 (2.1)	1.3 (0.4)	1.3 (0.5)	7.9 (1.9)	
Texas	1997–2017	0.98 (0.01)	99.0 (0.6)	2.0 (1.7)	0.4 (0.3)	0.9 (0.7)	1.1 (1.0)	
Colorado	1997–2017	0.95 (0.02)	97.6 (0.8)	4.0 (1.3)	0.9 (0.4)	1.0 (0.5)	3.7 (1.2)	
New Mexico	1997–2017	0.95 (0.03)	97.6 (1.5)	5.1 (3.9)	0.8 (0.6)	1.5 (1.0)	3.0 (2.3)	
East-LANID1997–2017	1997–2017	0.89 (0.01)	94.4 (0.6)	10.7 (1.1)	0.5 (0.1)	0.6 (0.1)	9.7 (0.9)	
East-LANID2012[a]	2012	0.49	74.4	49.4	1.9	3.7	33.5	
Eleven western states	1997–2017[b]	0.84 (0.06)	94.0 (2.9)	12.6 (7.5)	2.8 (0.6)	8.7 (5.4)	6.9 (6.7)	Ketchum et al. (2020)
NHPA[c]	2002 2015	0.92 0.86	97.0 96.3	0.9 17.1	3.7 0.9	11.3 4.9	0.3 3.5	Deines et al. (2017)
Arizona	2013, 2014, 2016	0.63 (0.06)	81.7 (3.1)	15.3 (5.6)	21.3 (1.6)	20.1 (1.7)	16.1 (4.9)	USGS (Brandt et al., 2021)
Washington	2015–2017	0.69 (0.04)	84.5 (2.0)	27.9 (0.1)	3.1 (3.8)	4.0 (4.7)	22.4 (0.8)	
Idaho	2011, 2015	0.87 (0.02)	93.6 (0.9)	10.4 (0.6)	2.5 (2.4)	2.7 (2.5)	9.6 (0.3)	
Colorado	2016	0.70	84.8	26.0	4.4	5.6	21.4	
Utah	2009	0.53	76.7	35.8	10.8	14.4	28.6	
Montana	2013	0.54	77.0	39.0	7.0	10.3	29.5	

Overall accuracy (OA), omission error (OE), and commission error (CE) are in percent. Note the validation area of the East was based on Fig. 1.
[a] LANID-2012 from Xie et al. (2019).
[b] The years without samples (1999) and having only one class of sample (i.e., 2004, 2005, 2012, 2015, and 2017) are not assessed
[c] The Northern High Plains Aquifer.

Compared to a previous approach for 30 m CONUS mapping (i.e. LANID 2012, Xie et al., 2019), the current strategy improved upon four aspects. First, we incorporated more training samples for the humid eastern states. While the previous 2012 LANID map used collected samples for only six states in the upper Midwest, our new training samples covered all remaining eastern states except for Arkansas, Mississippi, and Louisiana (where visual collection was difficult). Since classification in the eastern CONUS was sensitive to the quality and quantity of training data, expanding the sample database there greatly improved mapping accuracy (Table 3). Second, the new classification was more locally adaptive (i.e., stratified by county instead of ecoregion), which further reduced the impacts of spatial variations in cloud-free data availability and crop growth within ecoregions. Third, we applied an enhanced post-classification strategy based on the assumption that croplands equipped with irrigation systems tend to be used and irrigated frequently. This logic-based filtering was inapplicable in

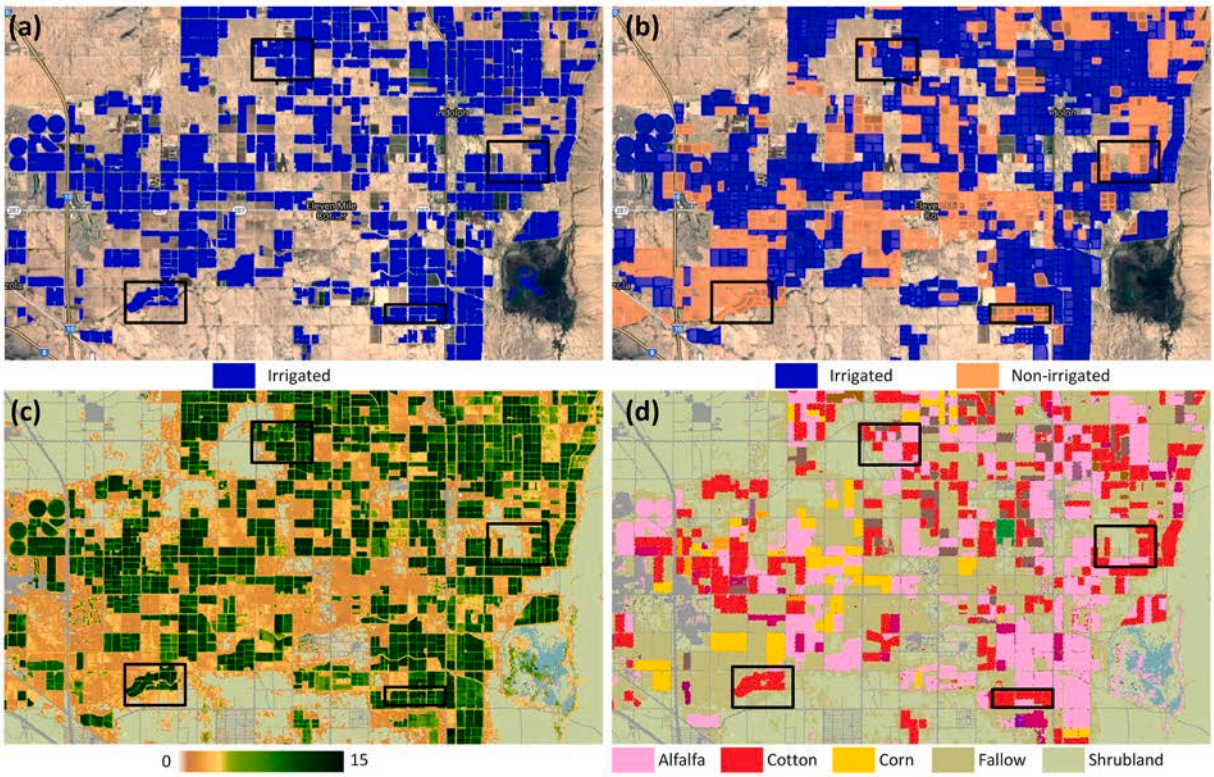

Fig. 10. Demonstration of inconsistency between 2013 LANID (a) and USGS-verified irrigated lands (b) in southern Arizona. The Landsat-derived GI (c) and USDA CDL (d) are used as comparisons. The black rectangles highlight mismatches between LANID and USGS data.

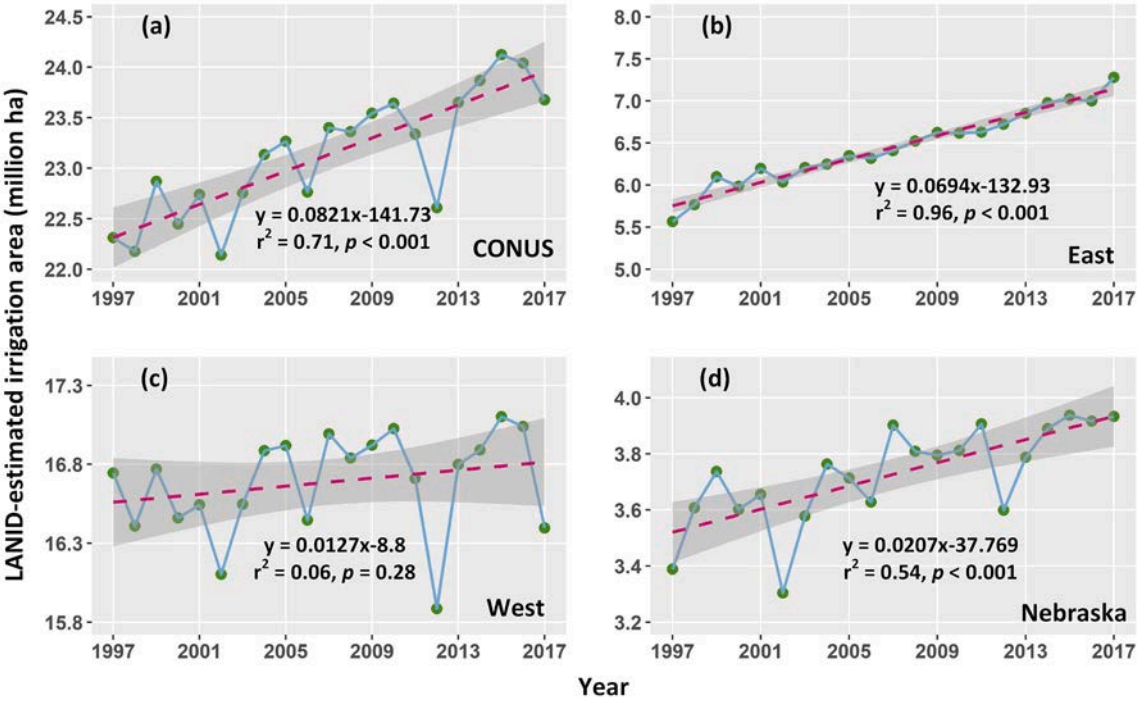

Fig. 11. LANID-derived irrigation trends during 1997–2017 at regional to country scales. The east-wide division is based on Fig. 1.

single-year irrigation mapping but proved to be instrumental in delineating and removing rainfed croplands falsely classified as irrigated in years with exceptional precipitation. Finally, we provided a more thorough nationwide accuracy assessment for the generated maps. Given the limited number of ground truth data, our previous 2012 LANID map was assessed in only six selected aquifers. This study, for the first time, presented a nationwide evaluation of irrigation maps based upon tens of thousands of test samples. In addition to the above improvements, our proposed method is temporally extendable, especially in the western CONUS where training sample collection can be

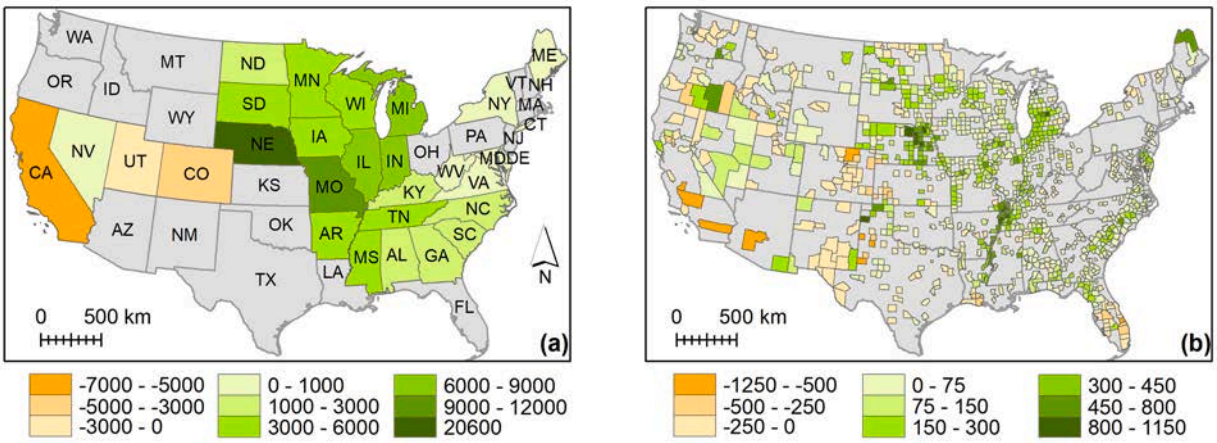

Fig. 12. LANID-derived trends during 1997–2017 at the state (a) and county (b) scale. States and counties with non-significant trends or having change rate smaller than one hectare are shown in gray.

automated for any year. For example, our threshold normalization strategy could be further applied to the years between 1985 and 1996 and after 2017 to map irrigated area for these years when Landsat images were available.

Both the qualitative and quantitative assessments showed good performance of LANID and indicated extensive improvements relative to other nationwide maps. Despite the advances, however, there remains several challenges and room for development in mapping irrigation across large spatial areas especially in humid regions. For example, LANID may have enlarged clusters of increasing trends for some humid states (e.g., Missouri, Michigan, North Dakota, and South Dakota) due to high confusion between irrigated and rainfed croplands. In addition, LANID, like most irrigation mapping approaches, used image composites as inputs to overcome the shortage of clear satellite observations. However, this strategy does not fully leverage the temporal profiles of crops captured by satellite observations, which could help identify essential time points when contrasts between irrigated and non-irrigated croplands are the greatest. This improved use of temporal information might be particularly valuable for irrigation mapping in the eastern CONUS where greenness contrasts between irrigated and non-irrigated croplands are weak in the maximum composites (Xie et al., 2019; Xu et al., 2019). While it might be impractical to characterize complete temporal profiles of crops for the years and regions with limited availability of clear Landsat imagery, integrating multi-source images can tackle the problem of data shortage in recent years, for example by combining Landsat and Sentinel images. In addition to increasing data availability of optical images, Sentinel sensors also provide time-series Synthetic Aperture Radar images that are particularly useful in irrigation detection in humid areas (Bazzi et al., 2020; Gao et al., 2018; Pageot et al., 2020). Another potential avenue to improve LANID in humid areas would be developing classification methods that account for specific crop types. Currently, most existing irrigation mapping approaches treat irrigated cropland as a single class, which can be problematic due to the phenological complexity of diverse crops. For instance, soybeans can have higher greenness than irrigated corn, potatoes, and dry beans, even under rainfed conditions (Deines et al., 2017; Xie et al., 2019). Therefore, crop-specific classifications would possibly enhance irrigation mapping accuracy, especially for areas with irrigated crops showing large ranges of peak greenness. Furthermore, the mapping accuracy of LANID in the Mississippi Alluvial Plain region remains uncertain due to the lack of reference samples and the difficulty of collecting ground truth data that confidently delineate irrigated vs. non-irrigated areas there.

Using the generated 21 annual irrigation maps, we identified several notable local to national irrigation trends for the country. In general, we found an overall increase of irrigated area across the CONUS, with most irrigation gains occurring in the Midwest, the Mississippi Alluvial Plain, and the East Coast regions. Whereas water scarcity hinders irrigation in the arid and semi-arid western U.S., conversion to irrigated agriculture in the eastern U.S. appears to be accelerating, carried on the backs of relatively more abundant and renewable surface water and groundwater (Mullen et al., 2009; Nocco et al., 2019). Groundwater depletion and recharge differences likely also account for many subregional variations of irrigation dynamics, such as the irrigation gains in the northern High Plains Aquifer vs. losses in the central and southern part of the aquifer and in the southern Central Valley Aquifer (DeAngelis et al., 2010; Deines et al., 2019; Scanlon et al., 2012).

Across the north central and eastern U.S., increasing demand for corn and soybeans for use as livestock feed, as well as expanded ethanol production in the Midwest, supported substantial recent cropland expansion (Lark et al., 2015; Smith et al., 2017) and may have similarly incentivized increased deployment and use of irrigation on existing croplands over the study period. Corn prices, for example, approximately doubled within the first decade of the 21st century (Brown and Pervez, 2014; Deines et al., 2017; Mullen et al., 2009) and these higher prices likely incentivized farmers to expand irrigation and seek better yields. Given that new croplands planted to corn and soybean have been found to yield on average 8–10% lower than existing crop extent (Lark et al., 2020), it may be that farmers preferentially irrigate existing croplands rather than convert natural lands in order to increase production and meet expanded demand.

Lastly, we found evidence of urbanization-induced loss of irrigated croplands in some areas, such as Phoenix, AZ, Boise, ID, and Fort Collins – Loveland, CO (Fig. S6). Here, rapid residential housing growth and urban development stimulated conversion of croplands to developed land, and in turn a reduction in irrigated extent. Given the competitions between urban and agricultural uses for land and water supplies – which is projected to be worsened by climate and population changes (Bren d'Amour et al., 2017; Elliott et al., 2014; Flörke et al., 2018), interfaces of conflict such as these may continue to grow in the future.

5. Conclusions

This study addressed the scientific need for field-relevant resolution, temporally frequent irrigation maps, which are crucial for the study and assessment of irrigation dynamics, water use, and associated socioeconomic and environmental consequences. To address this gap in knowledge and data, we created annual irrigation maps at 30 m resolution for the conterminous United States for the period of 1997–2017, implementing a new sampling strategy and logic-based post-classification methods. The produced LANID irrigation maps delivered reasonable performance, with overall accuracy above 90% for all regions and across all annual products. Based on the 21 years of data, we detected

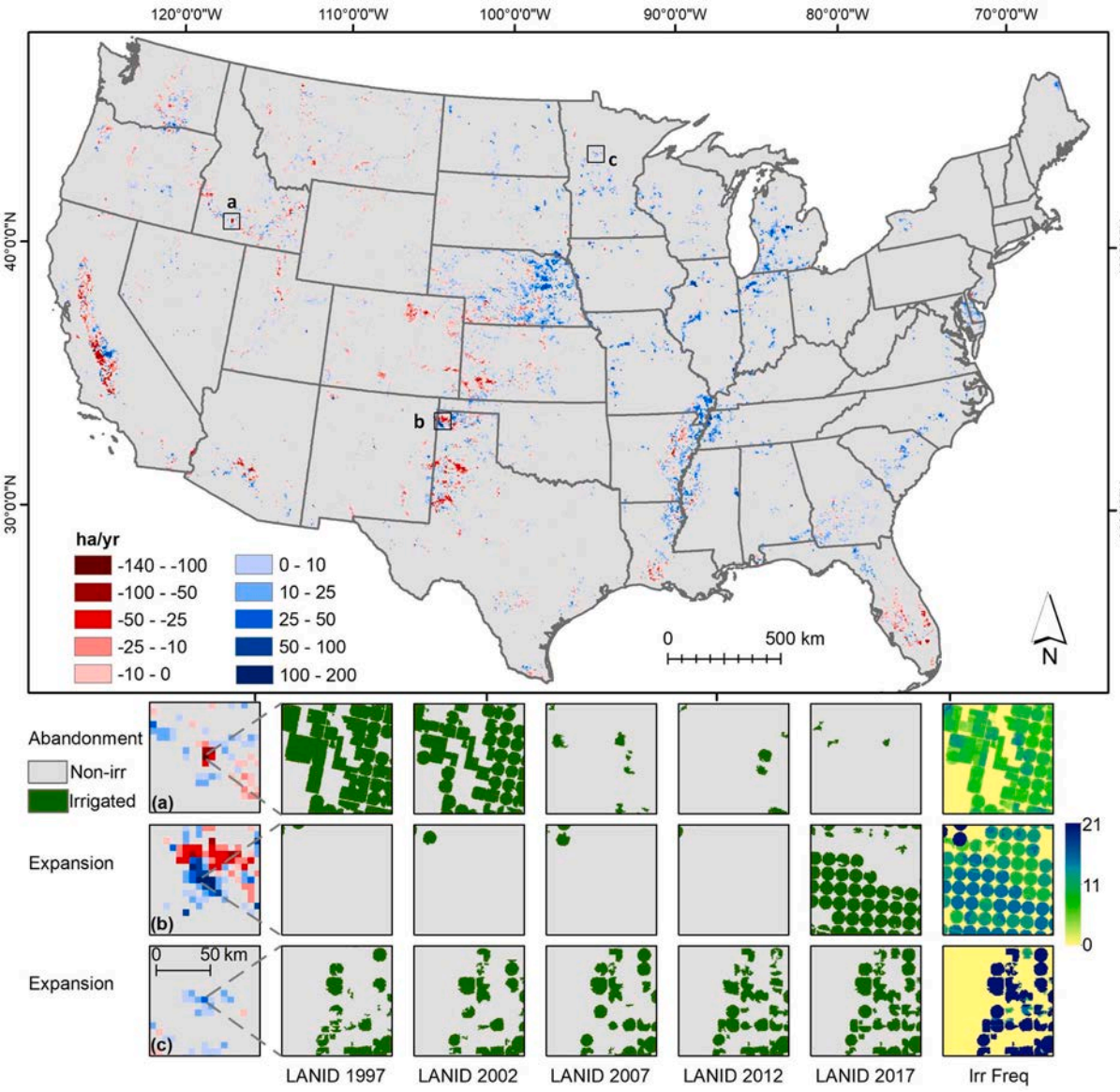

Fig. 13. LANID-derived spatially explicit irrigation trends during 1997–2017 at pixel scale. (a)-(c) show local views of irrigation changes in the central Idaho, northern Texas, and central Minnesota, respectively. The LANID insets highlihgt irrigation changes within a 6 km × 6 km grid. Rate of change is modeled using linear regression. Per-pixel irrigation trend is calculated per 6 km × 6 km grid with irrigated area aggregated from annual LANID maps. Grids without significant trends (p value of linear model >0.05) or having maximum irrigated area < 5% are shown in gray.

considerable increases in irrigated area especially in the Midwest, the Mississippi Alluvial Plain, and the East Coast. These areas of growth were counterbalanced, in part, by hotspots of decreasing irrigated extent in locations like the central and southern High Plains Aquifer, southern California Central Valley and Arizona, and southeastern Florida. The likely drivers behind these irrigation trends—such as water availability, crop prices, and urbanization—vary by region, as did the patterns of irrigation use and change themselves. The proposed mapping approach and associated data sets, which provide the most extensive and finest resolution characterization of U.S. irrigation to date, can help reveal a new level of understanding of irrigated land use and land-use change. Furthermore, the methods and data are temporally applicable to other years for continuous monitoring of irrigated area over CONUS and spatially extendable to other areas with similar climate and cropping landscapes. The LANID maps are available through Google Earth Engine (Asset Id: "users/xyhuwmir4/LANID/LANID_v1_rse") or requesting a copy from the corresponding author.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2021.112445.

Credit author statement

Yanhua Xie: conceptualization, methodology, formal analysis, writing - original draft, writing - review & editing; Tyler Lark: conceptualization, funding acquisition, writing - review & editing. All authors have reviewed and agreed on the published version of the manuscript.

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.

Acknowledgement

The authors would like to thank members of the USGS Water Budget

and Estimation Project for supplying of data, their feedback and ideas for mapping, and their insights regarding data use for water estimation. We also thank the anonymous reviewers and editors for their helpful comments. This research was based upon work supported by the United States Geological Survey under Award Number G19AC00080 and the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018409. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Bazzi, H., Baghdadi, N., Fayad, I., Zribi, M., Belhouchette, H., Demarez, V., 2020. Near real-time irrigation detection at plot scale using Sentinel-1 data. Remote Sens. 12, 1456. https://doi.org/10.3390/rs12091456.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Boryan, C., Yang, Z., Mueller, R., Craig, M., 2011. Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program. Geocarto Int. 26, 341–358.

Brandt, J.T., Caldwell, R.R., Haynes, J.V., Painter, J.A., Read, A.L., 2021. Verified Irrigated Agricultural Lands for the United States, 2002–17. U.S. Geological Survey data release. https://doi.org/10.5066/P9NAWU1U.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Bren d'Amour, C., Reitsma, F., Baiocchi, G., Barthel, S., Guneralp, B., Erb, K.H., Haberl, H., Creutzig, F., Seto, K.C., 2017. Future urban land expansion and implications for global croplands. Proc. Natl. Acad. Sci. U. S. A. 114, 8939–8944. https://doi.org/10.1073/pnas.1606036114.

Brown, J.F., Pervez, M.S., 2014. Merging remote sensing data and national agricultural statistics to model change in irrigated agriculture. Agric. Syst. 127, 28–40. https://doi.org/10.1016/j.agsy.2014.01.004.

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. https://doi.org/10.1016/j.rse.2017.10.030.

DeAngelis, A., Dominguez, F., Fan, Y., Robock, A., Kustu, M.D., Robinson, D., 2010. Evidence of enhanced precipitation due to irrigation over the Great Plains of the United States. J. Geophys. Res. 115. https://doi.org/10.1029/2010jd013892.

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, 9350–9360. https://doi.org/10.1002/2017gl074071.

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.

Dieter, C.A., Maupin, M.A., Caldwell, R.R., Harris, M.A., Ivahnenko, T.I., Lovelace, J.K., Barber, N.L., Linsey, K.S., 2018. Estimated Use of Water in the United States in 2015. In: US Geological Survey.

Drysdale, K.M., Hendricks, N.P., 2018. Adaptation to an irrigation water restriction imposed through local governance. J. Environ. Econ. Manag. 91, 150–165. https://doi.org/10.1016/j.jeem.2018.08.002.

Elliott, J., Deryng, D., Muller, C., Frieler, K., Konzmann, M., Gerten, D., Glotter, M., Florke, M., Wada, Y., Best, N., Eisner, S., Fekete, B.M., Folberth, C., Foster, I., Gosling, S.N., Haddeland, I., Khabarov, N., Ludwig, F., Masaki, Y., Olin, S., Rosenzweig, C., Ruane, A.C., Satoh, Y., Schmid, E., Stacke, T., Tang, Q., Wisser, D., 2014. Constraints and potentials of future irrigation water availability on agricultural production under climate change. Proc. Natl. Acad. Sci. U. S. A. 111, 3239–3244. https://doi.org/10.1073/pnas.1222474110.

Esri, 2010. USA Major Rivers. In: https://www.arcgis.com/home/item.html?id=290e4ab8a07f4d2c8392848d011add32.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., 2007. The shuttle radar topography mission. Rev. Geophys. 45.

Flörke, M., Schneider, C., McDonald, R.I., 2018. Water competition between cities and agriculture driven by climate change and urban growth. Nat. Sustain. 1, 51–58. https://doi.org/10.1038/s41893-017-0006-8.

Gao, Q., Zribi, M., Escorihuela, M., Baghdadi, N., Segui, P., 2018. Irrigation mapping using Sentinel-1 time series at field scale. Remote Sens. 10, 1495. https://doi.org/10.3390/rs10091495.

Gitelson, A.A., 2005. Remote estimation of canopy chlorophyll content in crops. Geophys. Res. Lett. 32 https://doi.org/10.1029/2005gl022688.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031.

Jackson, R.D., Idso, S., Reginato, R., Pinter Jr., P., 1981. Canopy temperature as a crop water stress indicator. Water Resour. Res. 17, 1133–1138.

Johnson, D.M., 2019. Using the Landsat archive to map crop cover history across the United States. Remote Sens. Environ. 232, 111286. https://doi.org/10.1016/j.rse.2019.111286.

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. https://doi.org/10.1016/j.jhydrol.2020.124905.

Ketchum, D., Jencso, K., Maneta, M.P., Melton, F., Jones, M.O., Huntington, J., 2020. IrrMapper: a machine learning approach for high resolution mapping of irrigated agriculture across the Western U.S. Remote Sens. 12, 2328. https://doi.org/10.3390/rs12142328.

Kraft, G.J., Stites, W., 2003. Nitrate impacts on groundwater from irrigated-vegetable systems in a humid north-central US sand plain. Agric. Ecosyst. Environ. 100, 63–74.

Lark, T.J., Meghan Salmon, J., Gibbs, H.K., 2015. Cropland expansion outpaces agricultural and biofuel policies in the United States. Environ. Res. Lett. 10, 044003 https://doi.org/10.1088/1748-9326/10/4/044003.

Lark, T.J., Mueller, R.M., Johnson, D.M., Gibbs, H.K., 2017. Measuring land-use and land-cover change using the U.S. department of agriculture's cropland data layer: cautions and recommendations. Int. J. Appl. Earth Obs. Geoinf. 62, 224–235. https://doi.org/10.1016/j.jag.2017.06.007.

Lark, T.J., Spawn, S.A., Bougie, M., Gibbs, H.K., 2020. Cropland expansion in the United States produces marginal yields at high costs to wildlife. Nat. Commun. 11, 4295. https://doi.org/10.1038/s41467-020-18045-z.

Lark, T.J., Schelly, I.H., Gibbs, H.K., 2021. Accuracy, Bias, and improvements in mapping crops and cropland across the United States using the USDA cropland data layer. Remote Sens. 13, 968. https://doi.org/10.3390/rs13050968.

Liu, H.Q., Huete, A., 1995. A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. IEEE Trans. Geosci. Remote Sens. 33, 457–465.

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L., Merchant, J.W., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. Int. J. Remote Sens. 21, 1303–1330.

Massey, R., Sankey, T.T., Congalton, R.G., Yadav, K., Thenkabail, P.S., Ozdogan, M., Sánchez Meador, A.J., 2017. MODIS phenology-derived, multi-year distribution of conterminous U.S. crop types. Remote Sens. Environ. 198, 490–503. https://doi.org/10.1016/j.rse.2017.06.033.

McDonald, R.I., Green, P., Balk, D., Fekete, B.M., Revenga, C., Todd, M., Montgomery, M., 2011. Urban growth, climate change, and freshwater availability. Proc. Natl. Acad. Sci. 108, 6312–6317.

Mullen, J.D., Yu, Y., Hoogenboom, G., 2009. Estimating the demand for irrigation water in a humid climate: a case study from the southeastern United States. Agric. Water Manag. 96, 1421–1428. https://doi.org/10.1016/j.agwat.2009.04.003.

Nocco, M.A., Smail, R.A., Kucharik, C.J., 2019. Observation of irrigation-induced climate change in the Midwest United States. Glob. Chang. Biol. 25, 3472–3484.

NRCS, 2016. SSURGO Web Soil Survey (USDA Nature Recources Conservation Service).

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, 3520–3537. https://doi.org/10.1016/j.rse.2008.04.010.

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 Sens. 12, 3044. https://doi.org/10.3390/rs12183044.

Pal, M., 2005. Random forest classifier for remote sensing classification. Int. J. Remote Sens. 26, 217–222.

Pervez, M.S., Brown, J.F., 2010. Mapping irrigated lands at 250-m scale by merging MODIS data and National Agricultural Statistics. Remote Sens. 2, 2388–2412. https://doi.org/10.3390/rs2102388.

Pervez, M.S., Budde, M., Rowland, J., 2014. Mapping irrigated areas in Afghanistan over the past decade using MODIS NDVI. Remote Sens. Environ. 149, 155–165. https://doi.org/10.1016/j.rse.2014.04.008.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J. Photogramm. Remote Sens. 67, 93–104. https://doi.org/10.1016/j.isprsjprs.2011.11.002.

Rosegrant, M.W., Ringler, C., Zhu, T., 2009. Water for agriculture: maintaining food security under growing scarcity. Annu. Rev. Environ. Resour. 34, 205–222. https://doi.org/10.1146/annurev.environ.030308.090351.

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. https://doi.org/10.1016/j.jag.2015.01.014.

Sampson, G.S., Hendricks, N.P., Taylor, M.R., 2019. Land market valuation of groundwater. Resour. Energy Econ. 58, 101120. https://doi.org/10.1016/j.reseneeco.2019.101120.

Scanlon, B.R., Faunt, C.C., Longuevergne, L., Reedy, R.C., Alley, W.M., McGuire, V.L., McMahon, P.B., 2012. Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley. Proc. Nat. Acad. Sci. 109, 9320–9325.

Seager, R., Ting, M., Li, C., Naik, N., Cook, B., Nakamura, J., Liu, H., 2012. Projections of declining surface-water availability for the southwestern United States. Nat. Clim. Chang. 3, 482–486. https://doi.org/10.1038/nclimate1787.

Seager, R., Lis, N., Feldman, J., Ting, M., Williams, A.P., Nakamura, J., Liu, H., Henderson, N., 2018. Whither the 100th Meridian? The once and future physical and human geography of America's arid–humid divide. Part I: the story so far. Earth Interact. 22, 1–22. https://doi.org/10.1175/ei-d-17-0011.1.

Seto, K.C., Guneralp, B., Hutyra, L.R., 2012. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. Proc. Natl. Acad. Sci. U. S. A. 109, 16083–16088. https://doi.org/10.1073/pnas.1211658109.

Shrestha, R.K., Cooperband, L.R., MacGuidwin, A.E., 2010. Strategies to reduce nitrate leaching into groundwater in potato grown in Sandy soils: Case study from north Central USA. Am. J. Potato Res. 87, 229–244. https://doi.org/10.1007/s12230-010-9131-x.

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. Discuss. 2, 1299–1327.

Siebert, S., Kummu, M., Porkka, M., Döll, P., Ramankutty, N., Scanlon, B.R., 2015. A global data set of the extent of irrigated land from 1900 to 2005. Hydrol. Earth Syst. Sci. 19, 1521–1545. https://doi.org/10.5194/hess-19-1521-2015.

Smith, T.M., Goodkind, A.L., Kim, T., Pelton, R.E., Suh, K., Schmitt, J., 2017. Subnational mobility and consumption-based environmental accounting of US corn in animal protein and ethanol supply chains. Proc. Natl. Acad. Sci. 114, E7891–E7899.

Tack, J., Barkley, A., Hendricks, N., 2017. Irrigation offsets wheat yield reductions from warming temperatures. Environ. Res. Lett. 12, 114027. https://doi.org/10.1088/1748-9326/aa8d27.

Teluguntla, P.G., Thenkabail, P.S., Xiong, J., Gumma, M.K., Giri, C., Milesi, C., Ozdogan, M., Congalton, R., Tilton, J., Sankey, T.T., 2015. Global Cropland Area Database (GCAD) Derived from Remote Sensing in Support of Food Security in the Twenty-First Century: Current Achievements and Future Possibilities. In: Taylor & Francis, Boca Raton, FL, USA.

Thenkabail, P.S., Biradar, C.M., Noojipady, P., Dheeravath, V., Li, Y., Velpuri, M., Gumma, M., Gangalakunta, O.R.P., Turral, H., Cai, X., 2009. Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium. Int. J. Remote Sens. 30, 3679–3733.

Thornton, P., Thornton, M., Mayer, B., Wilhelmi, N., Wei, Y., Devarakonda, R., Cook, R., 2016. Daymet: daily surface weather data on a 1-km grid for North America, Version 3. ORNL DAAC, Oak Ridge, Tennessee, USA. In: USDA-NASS, 2019. 2017 Census of Agriculture, Summary and State Data, Geographic Area Series, Part 51, AC-17-A-51. Washington D.C., USA.

Vashisht, B., Nigon, T., Mulla, D., Rosen, C., Xu, H., Twine, T., Jalota, S., 2015. Adaptation of water and nitrogen management to future climates for sustaining potato yield in Minnesota: field and simulation study. Agric. Water Manag. 152, 198–206.

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, 4626–4632. https://doi.org/10.1002/grl.50686.

Wardlow, B.D., Callahan, K., 2014. A multi-scale accuracy assessment of the MODIS irrigated agriculture data-set (MIrAD) for the state of Nebraska, USA. GISci. Rem. Sens. 51, 575–592.

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. https://doi.org/10.1016/j.isprsjprs.2019.07.005.

Xu, T., Deines, J., Kendall, A., Basso, B., Hyndman, D., 2019. Addressing challenges for mapping irrigated fields in subhumid temperate regions by integrating remote sensing and Hydroclimatic data. Remote Sens. 11, 370. https://doi.org/10.3390/rs11030370.

Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S.M., Case, A., Costello, C., Dewitz, J., Fry, J., Funk, M., Granneman, B., Liknes, G.C., Rigge, M., Xian, G., 2018. A new generation of the United States National Land Cover Database: requirements, research priorities, design, and implementation strategies. ISPRS J. Photogramm. Remote Sens. 146, 108–123. https://doi.org/10.1016/j.isprsjprs.2018.09.006.

Zektser, S., Loáiciga, H.A., Wolf, J., 2005. Environmental impacts of groundwater overdraft: selected case studies in the southwestern United States. Environ. Geol. 47, 396–404.

Zhu, Z., Woodcock, C.E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sens. Environ. 118, 83–94.

Zhu, Z., Wang, S., Woodcock, C.E., 2015. Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and sentinel 2 images. Remote Sens. Environ. 159, 269–277.